

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 13, 2003, 23:27:52 ; Search time 3327 Seconds
(without alignments)
5009.064 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 5241
Sequence: 1 MHILVHANYILLTLCPPRA.....MDPSAQISEALRLHMEAVM 1029

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues
Total number of hits satisfying chosen parameters: 33308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgnr_1/USPRO.spool/US09929769/Unat_07052003.161925-8044/app.query.fasta.1.1223
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-OUTFMT=lig=200 -NRM=SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US09929769.ecgn.1.1.2475 @unat.07052003.161925.8044 -NCPD=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-MARK_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gd_estl1:*
10: gd_estl2:*
11: gd_htc:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estlun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_oth:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query		ALIGNMENTS		Description	
No.	Score	Match	Length	DB	ID		
1	1543	29.4	981	9	AL518865	AL518865	AL518865
2	1461.5	27.9	959	14	B0071036	B0071036	AGENCOURT
3	1431.5	27.3	951	14	B0940921	B0940921	AGENCOURT
4	1411	26.9	971	13	BM548869	BM548869	AGENCOURT
5	1379	26.3	896	14	B0878711	B0878711	AGENCOURT
6	1374.5	26.2	896	14	B0686838	B0686838	AGENCOURT
7	1327.5	25.3	1054	14	BM810364	BM810364	AGENCOURT
8	1313	25.1	1098	13	BG761156	BG761156	AGENCOURT
9	1295	24.7	823	12	BG828224	BG828224	AGENCOURT
10	1292	24.7	800	12	BG825357	BG825357	AGENCOURT
11	1278.5	24.4	858	12	BG768219	BG768219	AGENCOURT
12	1278	24.4	897	14	BQ953580	BQ953580	AGENCOURT
13	1218	23.2	728	12	BG468191	BG468191	AGENCOURT
14	1217	23.2	899	12	BE792741	BE792741	AGENCOURT
15	1206	23.0	994	12	BE530701	BE530701	AGENCOURT
16	1205	23.0	730	9	AL042953	AL042953	DKFZP434L
17	1203.5	22.9	766	13	B1835118	B1835118	AGENCOURT
18	1199	22.8	729	13	B1554130	B1554130	AGENCOURT
19	1195	22.8	859	12	BF689825	BF689825	AGENCOURT
20	1190.5	22.7	840	12	BG765306	BG765306	AGENCOURT
21	1190	22.7	810	13	B1259297	B1259297	AGENCOURT
22	1182	22.6	1000	12	BG122941	BG122941	AGENCOURT
23	1172.5	22.4	970	14	BQ918808	BQ918808	AGENCOURT
24	1149	21.9	752	13	B1259197	B1259197	AGENCOURT
25	1137	21.9	872	12	BG424981	BG424981	AGENCOURT
26	1124	21.4	923	9	AL526487	AL526487	AGENCOURT
27	1118.5	21.3	1001	12	BG749177	BG749177	AGENCOURT
28	1111.5	21.2	1045	10	BE561214	BE561214	AGENCOURT
29	1111.5	21.2	706	13	BM011375	BM011375	AGENCOURT
30	1108	21.1	923	9	AL526445	AL526445	AGENCOURT
31	1104	21.1	833	12	BE729570	BE729570	AGENCOURT
32	1103.5	21.1	673	10	BE535502	BE535502	AGENCOURT
33	1100.5	21.0	736	12	BG822646	BG822646	AGENCOURT
34	1099.5	20.9	939	12	BF342499	BF342499	AGENCOURT
35	1093.5	20.8	799	13	B1091872	B1091872	AGENCOURT
36	1089	20.7	786	14	BQ215952	BQ215952	AGENCOURT
37	1087	20.7	786	14	BQ215952	BQ215952	AGENCOURT
38	1084.5	20.7	825	12	BG753046	BG753046	AGENCOURT
39	1074.5	20.5	825	12	BM018791	BM018791	AGENCOURT
40	1073	20.5	684	12	BG774322	BG774322	AGENCOURT
41	1068	20.4	748	13	BM050887	BM050887	AGENCOURT
42	1066	20.3	766	13	BE410337	BE410337	AGENCOURT
43	1059	20.2	1013	14	BQ715615	BQ715615	AGENCOURT
44	1057.5	20.2	1018	12	BG328409	BG328409	AGENCOURT
45	1057.5	20.2	1018	12	BG328409	BG328409	AGENCOURT

RESULT 1
AL518865
LOCUS
DEFINITION AL518865 LTI_NFL011.NBC1 Homo sapiens cDNA clone CSDBA011MY19 5
ACCESSION AL518865
VERSION AL518865.1 GI:12782358
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 961)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1. 981
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS04011YM19"
 /clone_1lb="LTI_NF1011_NBC1"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 173 a 328 c 303 g 164 t 13 others

ORIGIN

Alignment Scores:
 Pred. No.: 7, 97e-145
 Score: 1543.00
 Percent Similarity: 94.17%
 Best Local Similarity: 93.25%
 Query Match: 29,44%
 DB: 9
 Gaps: 0

US-09-929-769-7 (1-1029) x ALS18865 (1-981)

QY 614 SerValArgLysValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGlyAspSerVal 633
 |||||:::|||||
 Db 3 AGTGTGAGGAGGTGACGAGACCTGTGTAGCTGATCCAGAGTGGGAGACAGCGTG 62

QY 634 LeuGlyArgArgCysArgAspLeuLeuGlnLeuTyrrLeuGlnArgProGluLeuArg 653
 |||||:::|||||
 Db 63 CTGGGAGGAGGCTGCGAGACCTTCTCTCGCGGCTACTACAGCGCGGAGCKGCG 122

QY 654 ValProValProGluValLeuLeuHisSerGluGlyAlaIleSerSerValCysLys 673
 |||||:::|||||
 Db 123 GTGCCGCTGCTGAGGCTCTACTGACAGGAGGAGGCTGCCAGCGGACGCTGCGAG 182

QY 674 LeuAspGlyLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAspSerArgAla 693
 |||||:::|||||
 Db 183 CKGAGGAGACTCATCCACCCTTCATCCACCTCTCTGCGAGACACAGCAGCTCCGCGCG 242

QY 694 LeuGlnAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAlaValAlaHisPro 713
 |||||:::|||||
 Db 243 TTGGAGAACCGAGGAGGAGGAGGAGGCTGCGGAGCTGCGGCGGCGGAGCGCG 302

QY 714 LeuLeuLeuLeuArgHisLeuProMetIleAlaIleLeuLeuHisGlyArgThrHisLeu 733
 |||||:::|||||
 Db 303 CTGCTGTGCTGACGAGCCTGCGCATGAKGCGGCGCTCTGACAGGCGGCGCACCCACTC 362

QY 734 AsnPheGlnGluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeu 753
 |||||:::|||||
 Db 363 AACCTGACGAGAGTTCGCGGAGCAGAACACCTGAGCGCTCTGCGACAGCTGCGGCGCTG 422

QY 754 LeuGluLeuLeuGlnProHisValPheArgSerGlnHisGlnGlyAlaLeuTrpAspCys 773
 |||||:::|||||
 Db 423 CTGAGACGCTGACGAGCGCGCGCTGTCGAGGAGACAGGAGGCGGCTGCGGAGCTGC 482

QY 774 LeuLeuSerPheIleArgLeuLeuLeuAsnTyArgLysSerSerArgHisLeuAlaIle 793
 |||||:::|||||
 Db 483 CTCTGTCTTATCCGCGCTGCTGTAATTAACAGAGAGTCTCCGCGCATCTGCGCTGCC 542

QY 794 PheIleAsnLysPheValGlnPheIleHisLysTyrrIleThrTyraAsnAlaProAlaAla 813
 |||||:::|||||
 Db 543 TTGATCAACAAGAGTTTGCGACGATTATCATAGACATTACCTCATATGCCCGACGAGCC 602

QY 814 IleSerPheLeuGlnLysHisAlaAspProLeuHisAspLeuSerPheAspAsnSerAsp 833
 |||||:::|||||
 Db 603 ATCTCTTCTCTGCGAGAGACCGCGGACCCCTCCACGACCTGTCTTCGACACAGTGAC 662

QY 834 LeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeuProSerArgAspAspArgThr 853
 |||||:::|||||
 Db 663 CTGCTGATGCTGATATCTCTCTGAGGCTCAGCTTGCCTGCGACAGGAGGAGGCGCC 722

QY 854 AspArgLysLeuAspGlnGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 873
 |||||:::|||||
 Db 723 GACGAGAGCGCTGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 782

QY 874 SerValSerLeuPheThrProLeuThrAlaAlaGluMetAlaProTyrrMetLysArgLeu 893
 |||||:::|||||
 Db 783 AGCTCTCTCTCTGCTACCCCTGACCGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGG 842

QY 894 SerArgGlyGlnThrValGluAspLeuLeuGluValLeuSerAspIleAspGluMetSer 913
 |||||:::|||||
 Db 843 TCCGCGGAGCCAAACGCTGAGAGATCTCTGAGGCTTGAGTTACAAAGACGAGAGMTCC 902

QY 914 ArgArgArgProGluIleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSer 933
 |||||:::|||||
 Db 903 CGCGGAGAGCCCGAGACGCTGAGCTTCTCTGACCACTGCGAGCGGCTGAGAGAGCYCG 962

QY 934 AlaGluGluCysCysArg 939
 |||||:::|||||
 Db 963 GCCGAGGAGGTGTCGCG 980

RESULT 2
 BO071036 959 bp mRNA linear EST 02-APR-2002
 LOCUS AGENCOUPT_6853141 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5927385
 DEFINITION 5', mRNA sequence.
 ACCESSION BO071036
 VERSION BO071036.1 GI:19900082
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 959)
 NIH-MGC http://mgc.ncl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LICM2102 row: d column: 10
 High quality sequence stop: 705.

FEATURES
 source
 1. 959
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5927385"
 /clone_1lb="NIH_MGC_47"
 /tissue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Brain; Vector: pORF7; Site: 1: XhoI, Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in


```

Db      123 AAACCAAGAGCAGCCAGGAGCCATAGGCGGCGGATTCGGGGAGCCAG 182
Qy      182 LeuArgValLeuGlyProGluAspAlaLeuAGlyMePheLeuGlnIlePheProLeu 201
Db      183 CTCGGGGTGTGGGCGCTGAGGACGACTGCTGCGATGCTCTCCAGATTTTCCGCTC 242
Qy      202 SerProAspProArgTrpGlnSerSerSerProArgProValAlaLeuAlaLeuGln 221
Db      243 AGCCCGGAGCCCTGGTGGGAGAGCTCCAGTCCCGCCCGCTGGCGCTCCGCTGACAG 302
Qy      222 AlaLeuGlyGlnIleLeuAlaArgValValGlnGlySerProGluValProGlyIleThr 241
Db      303 GCCCTGGGAGGAGGCTGGCGCGCTGCTCAGGAGGAGCCCGAGGTGGCGGAGATCAGC 362
Qy      242 ValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyAlaLeuVal 261
Db      363 GTCCGCTCTCTGAGGCGCCCGCCAGCTCTGACTGAGTCCACACAGCGGGTCCCTGCTG 422
Qy      262 MetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuGlyGlnTyr 281
Db      423 ATGTCATGACCGCTAGCCACTTCTGCGCTGCGCCGCTGCGCGCAGCTCTGCAAGTAC 482
Qy      282 GluArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeuGln 301
Db      483 CAGCGCTGTGTGCCACAGGACCGGCTTCTCTGCTCTCTGAGGAGTCTCTGAC 542
Qy      302 MetLeuGlnTrpLeuAspSerProGlyValGlnGlyLeuProLeuArgAlaGlnLeuArg 321
Db      543 ATGCTGACAGTGGCTGACAGCCCTGGGGGTGAGGGGGCGGCGGAGACAGCTCAGG 602
Qy      322 MetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeuLeu 341
Db      603 ATGCTTGCCAGC----- 614
Qy      342 ArgLeuAlaGluAlaLeuAlaPheArgGlnAspLeuGluValValSerSerThrValArg 361
Db      615 ---CTGGCCGAGGCGCTGGCGCTTCCGTCAGGAGCTGGAGGTGCTCAGCTCCAGCTCCGT 671
Qy      362 AlaValIleAlaThrLeuArgSerGlyGlnGlnCysSerValGluProAspLeuIleSer 381
Db      672 GCCGTCATCCGACCCCTGAGAGTCTGGGGAGACAGTGCAGCTGGAGCGGAGCATGATCAGC 731
Qy      382 LysValLeuGlnIleLeu---IleGluValArgSerProHisLeuGlnIleLeuThr 401
Db      732 AAAGTCTCTCCAGGGGGCTGATCAAGGTGAGTCCCGCCACCTGGAGAGCTGCTGACTG 791
Qy      401 IapPheSerAlaThrAlaAspAlaAlaSerProPheProAlaCysLysPro-ValVal 420
Db      792 CATTTCTCTGTGCACGTGGGATGCTGCTCCCGCTTCCAGCTGTAAACCCNGTTGGG 851
Qy      421 ValValSerSerLeuLeuLeuGlnGlnGlnGluProLeuAlaGlyLysProGlyAla 440
Db      852 GGGGGAAACTCCCTGCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 911
Qy      441 -AspGlyGly-SerLeuGluAlaValArgLeuGly 451
Db      912 GGAGGGTGGCAACCTGGAAGCCCTGGCGTTGGG 946

```

RESULT 4
 LOCUS BMS48869 971 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOUNT_6543289 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5740963
 5', mRNA sequence.
 ACCESSION BMS48869
 VERSION BMS48869.1 GI:18783808
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 971)
 AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph. D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12756 row: m column: 16
High quality sequence stop: 636.
FEATURES
  source
    1..971
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:5740983"
    /clone_lib="NIH_MGC_88"
    /tissue_type="duodenal adenocarcinoma, cell line"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: small intestine; Vector: pCMV-SPORT6;
    Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally;
    oligo-dT primed. Average insert size 1.767 kb. Library
    enriched for full-length clones and constructed by Life
    Technologies. Note: this is a NIH_MGC library."
BASE COUNT 150 a 332 c 331 g 158 t
ORIGIN
Alignment Scores:
Pred. No.: 1,71e-131 Length: 971
Score: 1411.00 Matches: 277
Percent Similarity: 97.89% Conservative: 2
Best Local Similarity: 97.19% Mismatches: 4
Query Match: 26,928 Indels: 2
DB: Gaps: 0
US-09-929-769-7 (1-1029) x BMS48869 (1-971)
Qy      410 AlaserProPheProAlaCysLysProValValValSerSerLeuLeuGlnGln 429
Db      3 GCCTCCCGGTTTCCAGCCTGAAGCCCGCTGTGGTGTGAGCTCCCTGCTGACAGAG 62
Qy      430 GluGluProLeuAlaGlyGlyLysProGlyAlaAspGlyLysSerLeuGluAlaValArg 449
Db      63 GAGAGGCCCTGTGGTGGGGGAAAGCCGGGTGGGACGCTGGACCTGGAGGCCGTGGC 122
Qy      450 LeuGlyProSerSerGlyLeuLeuValAspTrpLeuGlnMetLeuAspProGluValVal 469
Db      123 CTGGGGCCCTCGTACAGCCCTCAAGTGTGACTGCTGGAATGCGAGACCCGAGGTGCTC 182
Qy      470 SerSerCysProAspLeuGlnLeuArgLeuLeuPheSerArgArgLysGlyGln 489
Db      183 AGCAGCTGCCCGCCAGCTCAGCTCAGCTCTCTTCCCGGAGAGAGAGAGAGAGAGAGAG 242
Qy      490 AlaGlnAlaProSerPheArgProTyrLeuLeuThrLeuPheThrHisGlnSerSerTrp 509
Db      243 GCCCAGGTGGCTCTGCTCCCTCTACCTCTGACCTCTTCAAGCATGATGATCAGCTGG 302
Qy      510 ProThrLeuHisGlnCysIleArgValLeuLeuGlnLysSerArgGlnGlnArgPheAsp 529
Db      303 CCCACACTGCACAGCTCATCCAGTCTCTGCGGCAAGAGCCGGGAACAGAGTTGAC 362
Qy      530 ProSerAlaSerLeuAspPheLeuTrpAlaCysIleHisValProArgIleTrpGlnGly 549
Db      363 CCTCTCTCTCTGAGACTCTCTCTGGGCTGCATCATGCTCTCTCTCATCTGACAGGG 422
Qy      550 ArgAspGlnArgThrProGlnLysArgArgGluGluLeuValLeuArgValGlnGlyPro 569
Db      423 CCGGACACGAGGACCCCGGAGAGCGGGGAGAGAGTGTGTGCTGCGGAGGCGCCG 482
Qy      570 GluLeuIleSerLeuValGluLeuIleLeuAlaGluAlaGluThrArgSerGlnAspGly 589

```


BASE COUNT	161 a	312 c	253 g	164 t	6 others
ORIGIN					
Alignment Scores:					
Pred. No.:	2,54e-128	Length:	896		
Score:	1379.00	Matches:	285		
Percent Similarity:	95.67%	Conservative:	2		
Best Local Similarity:	95.00%	Mismatches:	12		
Query Match:	26.31%	Indels:	3		
DB:	14	Gaps:	0		
US-09-929-769-7 (1-1029) x BQ878711 (1-896)					
QY 664	GIUGIYAlAAlaserSerSerValCysIysIeuAspJlyLeuIleHisArgPheIleThr	68:			
DB 1	GAAGGGGCTGGCAGAGGAGGCTGTGCAAGCTGGAGCGGACTATCCACGGCTTATCCG	60			
QY 684	LeuLeuAlaAspThrSerAspSerArgAlaLeuGIuaAsnArgIylAlaAspAlaSerMet	70:			
DB 61	CTCCCTGGGAGACACAGGAGCTCCGGGGCGTGGAGAACCCAGGGGCGGATGCCAGCATG	120:			
QY 704	AlaCysArgIylSerAlaAlaValAlaHisProIleuLeuLeuAlaGHisLeuProMetIle	72:			
DB 121	GCCTCCCGGAAGCTGGCGGCTGGCGCACCCGCTGTGCTGCTGACGACACTCCCATGTTC	180:			
QY 724	AlaAlaLeuLeuHisIGIYArgThrHisIeuAsnPheGInglIupPheArgIglIaHisHis	74:			
DB 181	GGGGGGCTCTGTCCACAGGGCGGACCCACTCAACTCCAGAGATTCCGGCAGACAGACAC	240:			
QY 744	IeuSerCysPheLeuHisValIleuGIylLeuLeuGIuLeuLeuGInProHisValPheArg	76:			
DB 241	CTGACCTCTCTTCACAGCTGTGGGCGCTGTGGAGCTGTGCACGCGCCAGCTGTCCGC	300:			
QY 764	SerGIuHisGInglYAlaLeuThrAspCysIeuLeuSerPheIleArgLeuLeuAsn	78:			
DB 301	AGCGAGCACAGAGGGGCGCTGTGGAGACTGCTTCTGTCTTCATCCGCTGTCTGTAT	360:			
QY 784	TyrArgIylSerSerArgHisIeuAlaAlaPheIleAsnIylPheValGInPheIleHis	80:			
DB 361	TACAGAGAGTCTCCCGGCACATCTGGCTCTTCATCAACAAAGTTGTGTCACATTGATCAT	420:			
QY 804	LysTyrIleThrTyrAsnAlaProIlaAlaIleSerPheLeuGInIylHisAlaAspPro	82:			
DB 421	AGTACATTTACCTACAAATGGCCCCACAGCCATCTCTTCTGTGCAGAACACGGCCGACCG	480:			
QY 824	IeuHisAspIeuSerPheAspAsnSerAlaPheValIleLeuIylSerLeuLeuAlaGly	84:			
DB 481	CTCCAGAGACTGTCTCTGTGCACACAGTGACCTGTGGATGTGAATCCCTCTGTGCAGGG	540:			
QY 844	IeuSerIeuProSerArgAspAspArgThrAspArgIylLeuAspGIuGIuGIuGIu	86:			
DB 541	CTCACCTGTCCCGACAGGAGGACACAGGACCGACCGACGGCTGTGAGAGAAAG--GGGAGAG	590:			
QY 864	GIuSerSerAlaGlySerIeuProIleuValSerValSerIeuPheThrProIleuThrAla	88:			
DB 600	GAGACTAGACCGGCTCTGTCTGTGCCCTGTGTACAGCTCTCCCTGTGTACCCCTGTACCGCG	650:			
QY 884	AlaGIuMetAlaProIylMetIylArgIeuSerArgIylGInIylThrValGIuAspIeuLeu	90:			
DB 660	GNCGAGATGGCCCCCTACATGMAAGCGCTTTTCCGCGGCCCAACAGCTGTGAGATCTGTCTG	710:			
QY 904	GIuValIeuSerAspIlaAspGIuMetSerArgIylArgIylProGInIleIeuSerPhePhe	92:			
DB 720	GAGGTTCTGAGTACATGACAGATGTTCGCCGGCGGAGAACCCGAGATCTGTAGCTTCTTC	770:			
QY 924	SerThrAsnLeuGInArgIeuMetSerSerAlaGIuGIuCysArgAsnLeuAlaPhe	94:			
DB 780	TGCAGCAANCTGCACGCTGATGATGACTCCCGCCGAGAGTGTGGCC--AACCTGCGCTTC	830:			
QY 944	SerIeu--AlaIeuArgSerMetGInAsnSerProSerIleuIleAlaAlaPheIeuPro	96:			
DB 839	AGCTGTGACCTGTGCTCCATGTCAAGAACGCCCCAGCATTCGAGCGGCTTTTGTCCCC	896			

RESULT 6
B0686838 896 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT 8171634 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5252041
DEFINITION 5', mRNA sequence.
ACCESSION B0686838
VERSION B0686838.1 GI:21812154
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2397 row: k column: 18
High quality sequence stop: 683.
Location/Qualifiers
1..896
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5252041"
/cui="NIH_MGC_110"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pONT7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(s). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."

BASE COUNT 157 a 307 c 274 g 158 t
ORIGIN

Alignment Scores:
Pred. No.: 7, 23e-128 Length: 896
Score: 1374.50 Matches: 281
Percent Similarity: 96.278 Conservative: 3
Best Local Similarity: 95.254 Mismatches: 7
Query Match: 26.234 Indels: 5
DB: 14 Gaps: 1

US-09-929-769-7 (1-1029) x B0686838 (1-896)

QY 562 LeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGluLeuIleLeuAlaGlu 581
|||||
DB 2 CTGCTGCTGGGGGTCCAGGGCCCGAGCTCATCAGCTGTGGAGTGCATCTGGCCGAG 61
|||||

QY 582 AlaGluThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArgLeu 601
|||||
DB 62 GCGGACACGGGAGGACAGGACGCGGACACGCGCTGCAGCTCATCCAGCGCGCTG 121
|||||

QY 602 ProLeuLeuLeuSerCysCysGlyAspAspGluSerValArgLysValThrGluHis 621
|||||
DB 122 CCCCTGCTGCTACGCTGCTCTGTGGGACGATGAGTGTCCAGAGGGTGACGGACAC 181
|||||

QY 622 LeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgCysArgAspLeu 641
|||||
DB 182 CTCTCAGGCTGATCCAGCAGAGTGGGAGACAGGCTGTGGGACGGCGCTCCGAGACCT 241
|||||

QY 642 LeuLeuGlnLeuTrpLeuGlnArgProGluLeuArgValProValProGluValLeuLeu 661
|||||

|||||
DB 242 CTCTGAGCTCTACCTACAGCGCGAGCTCGGGTGGCCCTGGCTGAGTCTACTG 301
|||||

QY 662 HisSerGluGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuIleHisArgPhe 681
|||||

DB 302 CACAGCGAAGGGCTGCCACAGCAGAGCTGTGCAAGCTGAGCGACTATCCACCGCTTC 361
|||||

QY 682 IleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGluAsnArgGlyAlaAspAla 701
|||||

DB 362 ATCAGCGCTCTCGGACACAGCAGACTCCCGGGCGTTGAGAAACGAGGGCGGATGCC 421
|||||

QY 702 SerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuAlaGlnHisLeuPro 721
|||||

DB 422 AGCATGGCTCCCGGAGAGCTGGGGTGGCCGACCCGCTGCTGTGCAGGCACTGGCC 481
|||||

QY 722 MetIleAlaLeuLeuHisGlyArgThrHisLeuAsnPheGlnIupPheArgGlnGln 741
|||||

DB 482 ATGATCGGGCGCTCTCGACGCGCCGACCCACTCACTTCCAGAGTTCCGGCAGACG 541
|||||

QY 742 AsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGluLeuGlnProHisVal 761
|||||

DB 542 AACCATCTGAGCTGCTCTCGACAGTGTGGCTGTGGAGCTGTGACGCGCAGCGACTG 601
|||||

QY 762 PheArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeuLeu 781
|||||

DB 602 TTCGGCAGGACGACACGAGGGGCGCTGTGGAGCTGCTTGTCTTCATCCGCTGTG 661
|||||

QY 782 LeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleLeuAsnLysPheValGlnPhe 801
|||||

DB 662 CTGATATACAGAAAGTCTCTCCGCACTGTGCTCTCATCAACAAGTTTGCACTTC 721
|||||

QY 802 IleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHisAla 821
|||||

DB 722 ATCATATAGTACATTACTCTCAATGCCCCAGCAGCATCTCTCTCGCA-AAAACAGGCC 780
|||||

QY 822 AspProLeuHisAspLeuSerPheAsp-AsnSerAspLeu-ValMetLeuLys- -SerL 840
|||||

DB 781 GACCCGCTCCAGCAGCTGCTTTAAACAACAGTGAAGCTGTGAGTCTTGAATCCCT 840
|||||

QY 840 euleuAlaGlyLeuSer-LeuProSerArgAspArg-852
|||||

DB 841 CCTTCGAGGGGTTCAGCTTGGCCCCGAGGAGGACCAACAG 879
|||||

RESULT 7
BM810364 1054 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT 6580583 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5453467
DEFINITION 5', mRNA sequence.
ACCESSION BM810364
VERSION BM810364.1 GI:19127187
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1054)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1947 row: a column: 20
High quality sequence stop: 632.
Location/Qualifiers
1..1054
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:5453467"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
BASE COUNT      172 a      358 c      336 g      186 t      2 others
ORIGIN
Alignment Scores:
Pred. No.:      5,47e-123      Length:      1054
Score:          1327.50      Matches:      274
Percent Similarity: 94.20%      Conservative: 2
Best Local Similarity: 93.52%      Mismatches: 10
Query Match:    25.33%      Indels:      7
DB:             14      Gaps:      1
US-09-929-769-7 (1-1029) x BM810364 (1-1054)
Oy 454 SerGlyLeuLeuValAspTrpLeuLumMetLeuAspProGluValSerSerCysPro 473
Db 1 TCAGGCTTCCTAGTGGAGCTGGGAAATGCTGACCCCGAGGTGTGACAGCTGCCCC 60
Oy 474 AspleuGlnLeuArgLeuLeuPheSerArgArgLySgLySgLyGlnAlaGlnAlaPro 493
Db 61 GACCTGCAGCTCAGCTGCTCTTCCCGGAGAGGGAAGGCAAGTCAGGCCAGCTGCC 120
Oy 494 SerPheArgProTrpLeuLeuThrLeuPheThrHisGlnSerSerTrpProThrLeuHis 513
Db 121 TCGTTCCTCCCTTACCTCTGACCTCTTCACGATCATGTCAGCTGGCCCACTGCAC 180
Oy 514 GlnCysIleArgValLeuLeuGlyLysSerArgGlnGlnArgPheAspProSerLaser 533
Db 181 CAGTGCATCCGAGTCTCTGCGGCAAGACCGGGAACAGGTTCGACCCCTCTGCTCT 240
Oy 534 LeuAspPheLeuTrpAlaCysIleHisValProArgIleTrpGlnGlyArgAspGlnArg 553
Db 241 CTGAGCTTCCTCTGGGCTGCATCATGTCCTGCATCTGCAGGGGGGAGCACAGCGC 300
Oy 554 ThrProGlnLysArgArgGlnGlnLeuValLeuArgValGlnGlyProGluLeuLieser 573
Db 301 ACCCGCGAAGCGCGGAGAGCTGTGTGCGGCTCCAGGCGCCGAGACTCATCACG 360
Oy 574 LeuValGluLeuIleLeuAlaGlnAlaGlnIleThrArgSerGlnAspGlyAspThrAlaAla 593
Db 361 CTGTGTGAGCTGATCTGTGGCGGAGGAGGAGCGGAGGCAAGAGGGGACACACCGCC 420
Oy 594 CysSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCysGlyAspAspGln 613
Db 421 TGCAGCTCATCCAGCGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Oy 614 SerValArgLysValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGlyAspSerVal 633
Db 481 AGTGTCCAGGAAGGTACGAGGACCTGTCAAGCTCATCCACAGTGGGGAACAGACGCTG 540
Oy 634 LeuGlyArgArgCysArgAspLeuLeuGlnLeuValTrpLeuGlnArgProGluLeuArg 653
Db 541 CTGGGCGAGCGCTGCGGAGACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Oy 653 GVal-ProValProGluValLeuLeuHisSer-GlnGlyAlaAlaSerSerSerValCys 672
Db 601 GGTGCGCCGTGCTGAGGCTTACTGACAGGNGAAGGGGCTGCCAGAGCGAGCTGCTC 660
Oy 673 LysLeuAspGlyLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAspSerArg 692
Db 661 AAGCTGACGAGCATCATCCACCGCTTTCACGCTTGTGCGAGACCGGACTCCCGG 720

```

```

Oy 693 AlaleuGluAsnArgGlyAlaAspAlaSerMetAlaCysArgLys-LeuAlaValAlaH 712
Db 721 GGTGTGAGAAAGCCGAGGGGCGGATGCGCTGGCGGAAACCTGGCGGTGGCGC 780
Oy 712 IsProLeuLeuLeuLeuArgHisLeuPheMet---IleAlaAlaLeuLeuHisGlyArgT 731
Db 781 ACCGCTGCTGTGCTGCTGAGGCGCCCTGCCATTCATGCGGGGCGCTCCGCGACGCCCA 840
Oy 731 hrHis-LeuAsnPhgGlnGluPheArg 739
Db 841 CCCACCTTAATTTTCAGGAATTCGCG 867
RESULT 8
Bg761156
LOCUS      831 bp      mRNA      linear      EST 15-MAY-2001
DEFINITION 602717492P1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4841134 5',
mRNA sequence.
ACCESSION  BG761156
VERSION     BG761156.1  GI:14071809
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 831)
REFERENCE   1
  AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE     Unpublished (1999)
  JOURNAL   Contact: Robert Strausberg, Ph.D.
            Email: c9apbs-remail.nih.gov
            Tissue Procurement: ATCC/DCTD/DRP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICM1673 row: 9 column: 23
            High quality sequence stop: 778.
            Location/Qualifiers
              1..831
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4841134"
                /clone_lib="NIH_MGC_49"
                /tissue_type="melanocytic melanoma, high MDR (cell line)"
                /lab_host="DH10B (phage-resistant)"
                /Note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:
                EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGCACGAG(G). Size-selected >500bp for average insert size
                1.8kb. Library constructed by Ling Hong in the Laboratory
                of Gerald M. Rubin (University of California, Berkeley)
                using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                II RT (Life Technologies). Note: this is a NIH_MGC
                library."
BASE COUNT      121 a      271 c      290 g      149 t
ORIGIN
Alignment Scores:
Pred. No.:      1.04e-121      Length:      831
Score:          1313.00      Matches:      267
Percent Similarity: 96.04%      Conservative: 0
Best Local Similarity: 96.04%      Mismatches: 10
Query Match:    25.05%      Indels:      3
DB:             12      Gaps:      1
US-09-929-769-7 (1-1029) x BG761156 (1-831)
Oy 332 ArgLeuSerAspValArgGlyLeuLeuArgLeuAlaGlnAlaLeuAlaPheArgGln 351
Db 2 AAGCTCAGTGTGCTGAGGGGCGCTCTGCGCTGGCGCGAGCCCTGCGCTTCCGTGAG 61

```


QY 649 gProgluLeuArgValProValProgluValLeuLeuHisSerGluGluAlaAlaSeSe 669
 Db 541 GCCGAGCTGGCGGGGCCGTGCTGAGAGTCTTACTGACAGCGAAGGGGCTGCCAGCAG 600
 QY 669 rSerValCysLysLeuAspGlyLeuIleHisArgPheIleIleHisLeuAlaAspThrSe 689
 Db 601 CAGGCTGCGACGCGGAGCGAGCTCATCCACCGCTTCATCCGCTTCGCGACACAG 660
 QY 689 rAspSerArgAlaLeuGluAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAl 709
 Db 661 CGACTCCCGGGCGCTGGAGAACCGAGGGCGGATGCCACATGGCTCCGAGACCTGGC 720
 QY 709 aValAlaHisProLeuLeuLeuLeu-ArgHisLeuProMetIleAlaAla-LeuLeuH 728
 Db 721 GGTGGCGACACCGCTGTGCTGCTCCAGGCGACCTGACCATGATGCGAGAGCTCTTGC 780
 QY 728 sGlyArgThrHisLeuAsnPheGlnGluPheArgGlnIleAsnHisLeuSerCysPheLe 748
 Db 781 ACGGACAGCCCACTCACTCCAGAGTCCGCGACAGAACCAACTGAGCTGTATCCT 840
 QY 748 uHisValLeuGly 752
 Db 841 GTACCTGCATGGG 853
 RESULT 10
 LOCUS Bg828224 823 bp mRNA linear EST 22-MAY-2001
 DEFINITION 602747827F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4906590 5',
 mRNA sequence.
 ACCESSION Bg828224
 VERSION Bg828224.1 GI:14175811
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 823)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LCM1809 row: 0 column: 07
 High quality sequence stop: 819.
 Location/Qualifiers
 1. 823
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4906590"
 /clone_1id="NIH_MGC_17"
 /tissue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: muscle; Vector: pOT7; Site_1: EcoRI;
 Site_2: XhoI; cDNA made by oligo-CT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GCGACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 133 a 275 c 270 g 145 t
 ORIGIN
 Alignment Scores: 1.35e-119 Length: 823
 Pred. NO.: 1292.00 Matches: 269
 Score:

Percent Similarity: 97.83% Conservative: 1
 Best Local Similarity: 97.46% Mismatches: 4
 Query Match: 24.65% Indels: 6
 DB: 12 Gaps: 0
 US-09-929-769-7 (1-1029) x Bg828224 (1-823)
 QY 530 ProSerAlaSerLeuAspPheLeuTrpAlaCysIleHisValProArgIleTrpGlnGly 549
 Db 3 CCCTTGCTCTCTGTGACTTCTCTGGGCTTCATCCATGTTCTCCGATCTGGCAGGG 62
 QY 550 ArgAspGlnArgThrProGlnLysArgArgGluLeuValLeuArgValGlnGlyPro 569
 Db 63 CCGGACACGACGCCACCCGACAGAGCGGAGAGTGTGTCTCGGGTCCA-GGCCCG 121
 QY 570 GluLeuIleSerLeuValGluLeuIleLeuAlaGluAlaGluThrArgSerGlnAspGly 589
 Db 122 GAGCTCATCAGCTGTGTGAGCTGATCTGTGGCGGAGGAGACCGAGCCGAGCCAGAGCGG 181
 QY 590 AspThrAlaAlaCysSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCys 609
 Db 182 GACACAGCGCGCTGACGCTCATCCAGGCGCGCTGCCCTGCTCAGCTGCTGT 241
 QY 610 GlyAspAspGluSerValArgLysValThrGluHisLeuSerGlyCysIleGlnGlnTrp 629
 Db 242 GGGGACGATGAGAGTGTGAGAGGAGTGACGAGACCTGTGAGCTGCATCCAGCAGTGG 301
 QY 630 GlyAspSerValLeuGlyAlaArgCysArgAspLeuLeuGlnLeuTrpLeuGlnArg 649
 Db 302 GGAGACAGCTGTGTGGAGGGGCTGTGGAGACCTTCTGTCTGACCTTACTTACAGCGG 361
 QY 650 ProGluLeuArgValProValProgluValLeuLeuHisSerGluGluAlaAlaSerSer 669
 Db 362 CCGGAGCTGGGGGCTGCCGTGCTGAGGCTTACTGACAGAGGAGGGGCTCCAGCAGC 421
 QY 670 SerValCysLysLeuAspGlyLeuIleHisArgPheIleIleHisLeuAlaAspThrSer 689
 Db 422 AGCCTGTCAAGCTGTGAGGAGCTCATCCACCCCTTCATCAGCCTTGTGGCAGCACAGC 481
 QY 690 AspSerArgAlaLeuGluAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAla 709
 Db 482 GACTCCCGGCTGTGAGAGCGAGGAGGGGATGCCAGCATGCTCCCGGAGAGCTGGCG 541
 QY 710 ValAlaHisProLeuLeuLeuLeuArgHisLeuProMetIleAlaAlaLeuLeuHisGly 729
 Db 542 GTGGGCGACCGGCTGTGCTGCTCAGGACACCTGCCATGATGGCGGCTCTTCGACAG 601
 QY 730 ArgThrHisLeuAsnPheGlnGluPheArgGlnIleAsnHisLeuSerCysPheLeuHis 749
 Db 602 CG-ACCCACCTTCACCTCCAGAGTTCGGCGAGCAGAACCACTGAGCTGTCTCTGCAC 660
 QY 750 ValLeuGlyLeuLeuGlnLeuLeuGlnProHisValPheArgSerGlnHisGlnGlyAla 769
 Db 661 GTGCTGGGCTGTGTGAGCTGTGCGAGCGGACGTGTTCGAGAGCAGCAGCAGGAGGCGG 720
 QY 770 LeuTrpAspCysLeu-LeuSerPheIleArgLeuLeuLeuAsnTrpArgLysSerSerArg 789
 Db 721 CTGTGGAGCTGCTTGTGTGCTTCTCATCCGCT-GTGGTGAATTCAGGAAGTCTCTCCG 779
 QY 789 gHis-LeuAlaAlaPheIleAsnLysPheValGlnPheIleHis 803
 Db 780 ACATCTGTGCTGCTTCATCAACAAGT-GTGCAGTTCATCAT 822
 RESULT 11
 LOCUS Bg825357 800 bp mRNA linear EST 22-MAY-2001
 DEFINITION 602747827F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4900615 5',
 mRNA sequence.
 ACCESSION Bg825357
 VERSION Bg825357.1 GI:14172944
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 800)
NIH-MGC <http://imgc.ncl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1794 row: f column: 08
High quality sequence stop: 800.

FEATURES

Location/Qualifiers
1..800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4900615"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pORF7; Site_1: EcoRI;
Site_2: XhoI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 126 a 272 c 266 g 136 t
ORIGIN

Alignment Scores:

Pred. No.: 2,96e-118 Length: 800
Score: 1278.50 Matches: 259
Percent Similarity: 96.64% Conservative: 0
Best Local Similarity: 96.64% Mismatches: 7
Query Match: 24.39% Indels: 4
DB: 12 Gaps: 1

US-09-929-769-7 (1-1029) x BG825357 (1-800)

QY 530 ProSerAlaSerLeuApsPheLeuTIPAlaCysIleHisValProArgIleTIPGlnGly 549
DB 3 CCCTGCGCCTCTGCGACTTCCTGGGCGCTGCATCCTGCTGCGATCTGGCAGGG 62
QY 550 ArgAspGlnArgThrProGlnLysArgArgGlnGlnLeuValLeuArgValGlnGlyPro 569
DB 63 CGGAGACCAAGGCAACCGCGGAGAGCGGAGAGCTGTGCTGCGGCTCA-GGCCCG 121
QY 570 GlnLeuIleSerLeuValGlnLeuIleLeuAlaGlnValGlnThrArgSerGlnAspGly 589
DB 122 GAGCTATCAGCTGTGGAGCTGATCTCTGGCGAGCGAGAGCGGACCGAGACGGG 181
QY 590 AspThrAlaAlaCysSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCys 609
DB 182 GACACAGCCGCTGCAGCTCATCCAGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCT 241
QY 610 GlnAspAspGlnSerValAlaGlyLysValThrGlnHisLeuSerGlyCysIleGlnGlnTTP 629
DB 242 GGGGACCATAGAGTGCAGGAAGGTGCAGGACACCTGTCAGGCTGCATCCAGCACTGC 301
QY 630 GlnAspSerValLeuGlnArgArgCysArgAspLeuLeuLeuGlnLeuTyrLeuGlnArg 649
DB 302 GGAAGACAGCTGCTGGCAGGCGCTGCGGAGCTTCTCTGACGCTTCACTACACGG 361
QY 650 ProGlnLeuArgValProValProGlnValLeuLeuHisSerGlnGlnAlaAlaSerSer 669
DB 362 CCGAGAGCTGGGGTCCCGCTGAGAGCTCTACTGCACAGCGAAGGGCTGCGCAGCAGC 421

QY 670 SerValCysLysLeuApsGlnLeuIleHisArgPheIleThrLeuLeuAlaAspThrSer 689
DB 422 AGGCTGCGAAGCTGGAGCGGACTCATCCACCGGTTATCATCGGCTCTGGGAGACACG 481
QY 690 AspSerArgAlaLeuGlnAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAla 709
DB 482 GACTCCCGGGCGTGGAGAACCGAGGGCGGATGCCAGCATGCGCTGCCGGAAGCTGGCG 541
QY 710 ValAlaHisProLeuLeuLeuLeuArgHisLeuProMetIleAlaLeuLeuHisGly 729
DB 542 GTGGCGACCCGCTGCTGCTGCTCAGGCACTGCCATGATCCGGCGCTCTGCGACGGC 601
QY 730 ArgThrHisLeuApsPheGlnGlnLeuPheArgGlnGlnAsnHisLeuSerCysPheLeuHis 749
DB 602 CGCACCACCTCAACTCCAGAGATTCCGCGACAGAACACCTGAGCTCTCTGCGAC 661
QY 750 ValLeuGlnLeuLeuGlnLeuLeuGlnProHisValPheArgSerGlnHisGlnGlyAl 769
DB 662 GTGCTGGCGCTGCTGGAGCTGCTGCAGCGCAGGTGTCCGACGACGACCAAGCGGGGC 721
QY 769 AleuTIPAspCysLeuLeuSerPheIleArgLeuLeuLeuAsnTyrArgLysSerSerArg 789
DB 722 GCTGTGGAGTGGCTTCTGCTTCATCCCGCTG-CCTGTAATTACAGAAAGTCTCCCGC 780
QY 789 GlnHisLeuAlaAlaPheIleAsn 796
DB 781 ATCT---GGCTGCTTCATCAAC 799

RESULT 12
BG768219 858 bp mRNA linear EST 15-MAY-2001
LOCUS 602744272F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4877441 5',
DEFINITION mRNA sequence.
ACCESSION BG768219
VERSION BG768219.1 GI:14078872
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 858)
NIH-MGC <http://imgc.ncl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCPD/DP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1761 row: p column: 18
High quality sequence stop: 750.

FEATURES

source

1..858
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4877441"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library. !"

```

BASE COUNT      128 a      281 c      294 g      155 t
ORIGIN

Alignment Scores:
Pred. No.:      3,78e-118      Length:      858
Score:          1278.00      Matches:      273
Percent Similarity: 95.52%      Conservative: 4
Best Local Similarity: 94.14%      Mismatches: 6
Query Match:      24.38%      Indels:      8
DB:              12      Gaps:      0

US-09-929-769-7 (1-1029) x BG768219 (1-858)

Qy      332 ArgLeuSerAspValArgGlyLeuLeuArgLeuAlaGluAlaLeuAlaPheArgGln 351
      |||||
Db      2 AGGCTCATGATGTCGAGGGGGGCTCTCGCCCTGGCCGAGAGCCCTCGCTCGTCAG 61

Qy      352 AspleuGluValSerSerThrValArgAlaValIleAlaThrLeuArgSerGlyGlu 371
      |||||
Db      62 GACCTGAGGTGTCAGCTCCACCGCTCCGTCATCGCCACCTGAGGCTGGGGAG 121

Qy      372 GlucysSerValGluProAspleuLeuSerLysValLeuGlnGlyLeuIleGluValArg 391
      |||||
Db      122 CAGTGCACGCTGAGACCGGACCTGATCAGCAAGTCTCCAGGGGCTGATCAGGTGAGG 181

Qy      392 SerProHISLeuGluGluLeuLeuThrAlaPhePheSerAlaThrAlaAspAlaLaser 411
      |||||
Db      182 TCCCCCACCCTGGAGAGACTGCTGACTGCTCTCTCTGACACGCGGATGCTGCTCC 241

Qy      412 ProPheProAlaCysLysProValValValSerSerLeuLeuGlnGluGln 431
      |||||
Db      242 CCGTTTCAGACCTGTAAACCCGTTGTGTGGTGGAGCTCCCTGCTCGACAGGAGAGAG 301

Qy      432 ProLeuAlaGlyGlyLysProGlyAlaAspGlyGlySerLeuGluAlaValArgLeuGly 451
      |||||
Db      302 CCCCCTGGGTGGGGAGACCGGGTGGGAGCGGTGGACCTGGAGCCGCTGGCGTGGGG 361

Qy      452 ProSerSerGlyLeuLeuValAspThrLeuGluMetLeuAspProGluValValSerSer 471
      |||||
Db      362 CCGCTGTAGGCTCTCTGAGTGTGAGTGTGAGAAATCTGAGACCCCGAGGTGCTGACGAC 421

Qy      472 CysProAspleuGlnLeuArgLeuLeuPheSerArgArgLysGlyLysGlyGlnAlaGln 491
      |||||
Db      422 TGCCCCGACCTGACACTGAGGCTGCTCTCTCCCGAGAGAGGCAAGGTGACGCCGAG 481

Qy      492 ValProSerPheArgProGlyLeuLeuThrLeuPheThrHisGlnSerSerThrProThr 511
      |||||
Db      482 GTGCCCTGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 541

Qy      512 LeuHISGlnCysIleArgValLeuLeuGlnGlyLysSerArgGlnGlnArgPheAspProSer 531
      |||||
Db      542 CTGCAACGAGTGTGATCCGAGTCTCTGAGGCAAGACCCGAGGAAAGAGTTTCACCCCTCT 601

Qy      532 AlaSerLeuAspPheLeuThrAlaCysIleHisValProArgLysLeuPheGln-GlyArgAs 551
      |||||
Db      602 GCGTCTGTGAGCTCTCTCTGAGGCTGAGTCCATGTTCTCTCCGATCTGAGAGGCGCGGA 661

Qy      551 PGLInArgThrProGlnLysArgArgGlu-GluLeuVal-LeuArgValGlnGly-ProGln 570
      |||||
Db      662 CCAGGACCCCGAAGAGCGGAGGAGGAGCTGTTGCTGCGGAGTCCAGGAGCCCGCA 721

Qy      570 uLeuIleSerLeu-ValGluLeuIleLeuAlaGluAlaGlu-ThrArg-SerGlnAspGln 589
      |||||
Db      722 GCTCATCAGCTGGGTGAGTGTGATCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 781

Qy      589 yAspThrAlaAlaCysSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCysCys 609
      |||||
Db      782 GCACACAGCCCGCATG-AGCCATTCACAGCGCGGTGCCCGCTGTTCAGTTGCTGCTG 840

Qy      609 sGlyAspAspGlnSer 614
      |||||
Db      841 TGGAGACGATTCACAG 856

```

```

RESULT 13
B0953580
LOCUS
DEFINITION
AGENCOURT 8882243 lupski.sciatric.netve Homo sapiens CDNA clone
IMAGE:6197684 5', mRNA sequence.
ACCESSION
B0953580
VERSION
B0953580.1 GI:22369058
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 897)
AUTHORS
NIH-MGC http://mgc.ncl.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LNL13608 row: b column: 21
High quality sequence stop: 576.
FEATURES
Source
1..897
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6197684"
/clone_lib="lupski.sciatric.netve"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/note="Vector: pcMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
directionally cloned using the following adaptors:
5'-TCGACCCAGCGGCTCG-3' and
5'-GACATGTTCTAGATCCGAGCGGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT      145 a      316 c      295 g      141 t
ORIGIN

Alignment Scores:
Pred. No.:      4.75e-112      Length:      897
Score:          1218.00      Matches:      244
Percent Similarity: 93.21%      Conservative: 3
Best Local Similarity: 92.08%      Mismatches: 13
Query Match:      23.24%      Indels:      5
DB:              14      Gaps:      3

US-09-929-769-7 (1-1029) x B0953580 (1-897)

Qy      465 AspProGluValValSerSerCysProAspleuGlnLeuArgLeuLeuPheSerArg 484
      |||||
Db      3 GACCCGAGGTGTGTCAGAGCTGCCCGACCTGAGCTGCTCTCTCTCTCTCTCTCTCT 62

Qy      485 LysGlyLysGlyGlnAlaGlnValProSerPheArgProGlyLeuLeuThrLeuPheThr 504
      |||||
Db      63 AAGGGCAAGGTGACAGCCAGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122

Qy      505 HisGlnSerSerThrProThrThrLeuHisGlnCysIleArgValLeuLeuGlnLysSerArg 524
      |||||
Db      123 CATCAGTCCAGCTGCGCACCTGACCATGTCATCCAGTCTCTCTGAGGAGGCGG 182

Qy      525 GluGlnArgPheAspProSerAlaSerLeuAspPheLeuThrAlaCysIleHisValPro 544

```

```

|||||
Db 183 GAACACAGGTTGACCCCTGCTCTGAGCTTCTGCGCCATGCTCC 242
Oy 545 ArgIleTgInglYArGspGlnArGThrProGlnLysArGArgGlnGluLeuValLeu 564
Db 243 CCGATCTGGCAGGGGGGAGCAGCCAGCCCGCAGAACGGGGAGAGAGCTGCTG 302
Oy 565 ArgValGInglYProGlnLeuLeuLeuSerLeuValGluLeuLeuLeuLeuLeu 584
Db 303 CCGGCTCCAGGGCCCGGAGCTCATCAGCTGAGTGTGAGCTGCTGCGGAGAGAG 362
Oy 585 ArgSerGlnAspGlyAspThrAlaAlaCysSerLeuLeuLeuLeuLeuLeuLeu 604
Db 363 CGAGCCAGAGAGGGGAGCAGAGCCCTGAGCTCATCCAGGCGCGCTGCTGCTG 422
Oy 605 LeuSerCysCysCysGlyAspAspGlnSerValArgValThrGlnHisLeuSerGly 624
Db 423 CTCAGCTGCTGCTGCTGGGAGCATGAGTGTGAGAGTGTGAGAGCAGCCTGTACG 482
Oy 625 CysIleGlnGlnTgPglYAspSerValLeuGlyArgArgCysArgAspLeuLeuGln 644
Db 483 TGCATCCAGCAGTGGGAGACAGCCTGTGCGCAGCGCTGCGGAGACCTTCTCTG 542
Oy 645 LeuTyrLeuGlnArGProGlnLeuLeuArgValProValProGlnValLeuHisSerGln 664
Db 543 CTCCTACTTACAGGCGGCGGAGCTGCGGCTGCGCTGAGTCTTCTACTGACAGCGAA 602
Oy 665 GlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuLeuHisArgPheLeuLeu 684
Db 603 GGGGCTGCCAGAGCAGCGCTGTCAAGCTGAGAGGAGCTATCCAGCGCTTCTCATCAG 662
Oy 685 LeuAlaAspThrSerAspSerArgAlaLeu---GluAsnArgGlyAlaAspAlaSerMet 703
Db 663 CTTCGCGAGACAGCAGCCAGCCCGGCGTGGCAGAACCGAGGGCGGATGCGCACCATG 722
Oy 704 Ala-CysArgLysLeuAlaValAla-----HisProLeuLeuLeuLeuArgHisLeu-- 720
Db 723 GCCCTCCCGGAAACTGCGGGGGGGGGGGCCCCCGCTGCTTCTGCTGCTGCTGCTG 782
Oy 721 -PrometIleAla 724
Db 783 CCCATGATCGCG 795

RESULT 14
LOCUS BG468191 728 bp mRNA linear EST 21-MAR-2001.
DEFINITION 602509766F1 NIH_MGC_15 Homo sapiens CDNA clone IMAGE:464040 5',
ACCESSION BG468191
VERSION BG468191.1 GI:13400461
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 728)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CMI415 row: c column: 17
High quality sequence stop: 724.
Location/Qualifiers
1..728
/organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:464040"
/clone_1ib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: colon; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 122 a 244 c 114 t
ORIGIN
Alignment Scores:
pred. No.: 4 1e-112 Length: 728
Score: 1217.00 Matches: 228
Percent Similarity: 99.17% Conservative: 2
Best Local Similarity: 98.35% Mismatches: 2
Query Match: 23.22% Indels: 1
DB: 12 Gaps: 0
US-09-929-769-7 (1-1029) x BG468191 (1-728)
Oy 525 GluGlnArGpheaSPProSerAlaSerLeuAspPheLeuTrpAlaCysIleHisValPro 544
Db 2 GAACAGAGGTTGACCCCTGCTGCTGAGTCTTCTGCGGCTGAGCTCATCCATGTTCT 61
Oy 545 ArgIleTgInglYArGspGlnArGThrProGlnLysArGArgGlnGluLeuValLeu 564
Db 62 CGCATCTGGAGGGGGGAGCAGCGCACCCCGCAGAGGGGGAGAGAGCTGTGCTG 121
Oy 565 ArgValGInglYProGlnLeuLeuLeuSerLeuValGluLeuLeuLeuLeuLeuLeu 584
Db 122 CCGGCTCCAGGGCCCGGAGCTCATCAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 181
Oy 585 ArgSerGlnAspGlyAspThrAlaAlaCysSerLeuLeuLeuLeuLeuLeuLeuLeu 604
Db 182 CGAGCCAGAGCAGCGGAGCAGCAGCGCTGAGCTCATCCAGCGCCCGCTGCTGCTG 241
Oy 605 LeuSerCysCysCysGlyAspAspGlnSerValArgValThrGlnHisLeuSerGly 624
Db 242 CTCAGCTGCTGCTGCTGGGAGCATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 301
Oy 625 CysIleGlnGlnTgPglYAspSerValLeuGlyArgArgCysArgAspLeuLeuGln 644
Db 302 TGCATCCAGCAGTGGGAGAGCAGCTGTGAGAGCGGTGCGGAGACCTTCTCTGAG 361
Oy 645 LeuTyrLeuGlnArGProGlnLeuLeuArgValProValProGlnValLeuHisSerGln 664
Db 362 CTCCTACTTACAGGCGGCGGAGCTGCGGCTGCGCTGAGTCTTCTACTGACAGCGAA 421
Oy 665 GlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuLeuHisArgPheLeuLeu 684
Db 422 GGGGCTGCCAGAGCAGCGCTGTCAAGCTGAGAGCTATCCACCGCTTCTCATCAGCTC 481
Oy 685 LeuAlaAspThrSerAspSerArgAlaLeuGlnAspArgGlyAlaAspAlaSerMetAla 704
Db 482 CTTCGCGAGACAGCAGCAGCTCCGGGCTTGGAGAACCGAGGGCGGATGCGCACATGGCC 541
Oy 705 CysArgLysLeuAlaValAlaHisProLeuLeuLeuLeuLeuLeuLeuLeuLeu 724
Db 542 TGCAGGAGAGTGGCGGTGGGAGCAGCGCTGCTGCTGAGCAGCAGTGTGCGCCATGATCGCG 601
Oy 725 AlaLeuLeuHisGlnYArgThrHisLeuAsnPheGlnGlnPheArgGlnGlnHisLeu 744
Db 602 GCGCTCTGACAGGGCC -AACCACTCACTTCAGAGAGTTCGGAGAGCAGAACCACTG 660
Oy 745 SerCysPheLeuHisValLeuGlnLeuLeuGlnLeuLeuGlnProHisValPheArgSer 764
Db 661 AGCTGCTTCTGACAGTGTGCGCTGTGAGCTGTGAGCTGTGAGCGGAGAACGTGTACCGCAGC 720

```


QY 765 GIuHIS 766
 Db 721 GAGCAC 726
 RESULT 15
 BE792741
 LOCUS 601585088F1 NIH_MGC_7 Homo sapiens CDNA clone IMAGE:393692 5',
 DEFINITION mRNA sequence.
 ACCESSION BE792741
 VERSION BE792741.1 GI:10213939
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 899)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at: Image.lnl.gov
 Plate: LCM/88 row: 0 column: 21
 High quality sequence stop: 771.
 Location/Qualifiers
 1..899
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="393692"
 /clone_id="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pORF7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 159 a 290 c 294 g 156 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,79e-111 Length: 899
 Score: 1206.00 Matches: 260
 Percent Similarity: 90.69% Conservative: 3
 Best Local Similarity: 89.66% Mismatches: 24
 Query Match: 23.01% Indels: 8
 DB: 12 Gaps: 1
 US-09-929-769-7 (1-1029) x BE792741 (1-899)
 QY 534 LeuAspPheLeuTrpAlaCysIleHisValProArgIleTrpGlnGlyArgAspGlnArg 553
 Db 3 CTGAGACTTCCTCTGGGCTTCATCATGTCTCGCATCTGGACGAGCGGACCGCC 62
 QY 554 ThrProGlnArgArgGlnGlnGlnValLeuArgValGlnGlnGlyProGlnLeuLeuSer 573
 Db 63 ACCCGCGAGAGCGGCGGAGAGCTGGCTGGCTGCA-GGCGCGAGCTCATTCAGC 121
 QY 574 LeuValGlnLeuIleLeuAlaGlnAlaGlnThrArgSerGlnAspGlnAspThrAlaAla 593
 Db 122 CTGGTGGAGCTGATCTCTGGCGGAGCGAGACCGGAGCCAGAGCGGGGACACAGCCGCC 181

QY 594 CysSerLeuIleGlnAlaArgLeuProLeuLeuSerCysCysGlyAspAspGln 613
 Db 182 TGCAGCTTCATCCAGCGCGGCTGCCCTGCTGCTAGCTGCTGGGGACATGAG 241
 QY 614 SerValArgGlyValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGlyAspSerVal 633
 Db 242 AGTTCAGGAAGGTACGAGGACCTGTACAGCTCATTCACAGCATGGGGAGACAGCTG 301
 QY 634 LeuGlyArgGlyCysArgAspLeuLeuGlnGlnLeuTrpGlnArgProGlnLeuArg 653
 Db 302 CTGGGAGAGCGCTGCCGAGACCTTCTCTGCAAGCTTCACTTACAGCGCGGAGCTGGCG 361
 QY 654 ValProValProGlnValLeuLeuHisSerGlnGlyAlaAlaSerSerSerValCysLys 673
 Db 362 GTGCCCGTGGCGTGGAGCTCTACTGACAGCGAGAGGGGCTGCCAGAGCGCTCTGCAG 421
 QY 674 LeuAspGlyLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAspSerArgAla 693
 Db 422 CTGGAGGAGCTCATCCACCGCTTCATCAGCTCTTGGCGGACACCGAGACTCCCGGCG 481
 QY 694 LeuGlnAsnArgGlyAlaAspAlaSerMetAlaCysArgGlyLeuAlaValAlaHisPro 713
 Db 482 TTGGAGAACCGAGGGGCGGATGCCAGCATGGCTCCGAGAACTGGCGGTGGCGACCCG 541
 QY 714 LeuLeuLeuLeuArgHisLeuProMetIleAlaAlaLeuLeuHisGlyArgThrHisLeu 733
 Db 542 CTGCTGCTGCTCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598
 QY 734 AsnPheGlnGluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeu 753
 Db 599 AACTTCCAGGAGTTCGGCGAGCAGAACCACTGAGCTGCTGCTGCTGCTGCTGCTGCTG 658
 QY 754 LeuGlnLeuLeuGlnProHisValPheArg-SerGlnHisGlnGlyAlaLeuTrpAspCys 773
 Db 659 CTGGAGCTGCTGACCGCGACGCTGTCCGAGCGGACGACGAGGGGCGCTGTGGAGCTG 718
 QY 773 SLeuLeuSerPheIleArgLeuLeuLeuAsnTyrArg-LysSerSerArgHisLeuAla 793
 Db 719 CTT-CTGTCTTCATCCCGCTGCTGCTG-ATTACAGGGAGAGTCTCCCGCATCTGGGCTG 776
 QY 793 IapHeIleAsnLysPheValGlnPheIleHisLysTyrIleThrTyrAsnAlaProAla 813
 Db 777 CTTTCATCAACAGTTTGTGCAAGTTCAATCCATTAAG--GTAAATATCCCAACAGGCGCA 833
 QY 813 IalleSerPheLeuGlnLysHisAla.821
 Db 834 GAGCATTTCTTCTTGAAGAGACGCG 859

Search completed: May 14, 2003, 02:29:53
 Job time : 3359 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:41:03 ; Search time 28 Seconds
(without alignments)
3381.945 Million cell updates/sec

Title: US-09-929-769-7

Perfect score: 5241

Sequence: 1 MHILVHAMVILTLTGPPRA.....MDPSAQISEALRIHMEAYM 1029

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 segs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTOS_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5241	100.0	1029	US-10-033-245-22	Sequence 22, Appl
2	5241	100.0	1029	US-10-033-223-22	Sequence 22, Appl
3	5241	100.0	1029	US-10-033-167-22	Sequence 22, Appl
4	5241	100.0	1029	US-10-033-244-22	Sequence 22, Appl
5	5241	100.0	1029	US-10-033-435-22	Sequence 22, Appl
6	5241	100.0	1029	US-10-032-990-22	Sequence 22, Appl
7	5241	100.0	1029	US-09-929-769-7	Sequence 7, Appl
8	5241	100.0	1029	US-10-032-996-22	Sequence 22, Appl
9	5241	100.0	1029	US-10-033-396-22	Sequence 22, Appl
10	5241	100.0	1029	US-10-033-246-22	Sequence 22, Appl
11	5241	100.0	1029	US-10-033-301-22	Sequence 22, Appl
12	5241	100.0	1029	US-10-033-326-22	Sequence 22, Appl
13	2055	39.2	428	US-09-984-271-235	Sequence 235, App
14	1665	31.8	308	US-09-984-271-148	Sequence 148, App
15	530.5	10.1	161	US-09-764-891-4155	Sequence 4155, Ap
16	156	3.0	2472	US-09-815-242-5064	Sequence 5064, Ap
17	135	2.6	1987	US-10-132-382-6	Sequence 6, Appl
18	135	2.6	2013	US-10-132-382-2	Sequence 2, Appl
19	135	2.6	2014	US-10-132-382-8	Sequence 8, Appl

20	135	2.6	2040	US-10-132-382-4	Sequence 4, Appl
21	131.5	2.5	1257	US-10-109-324-2	Sequence 2, Appl
22	127.5	2.4	3353	US-09-888-615-64	Sequence 64, Appl
23	125	2.4	2383	US-10-082-830-260	Sequence 260, App
24	123	2.3	693	US-10-029-217A-4	Sequence 4, Appl
25	122.5	2.3	4999	US-09-976-059-14	Sequence 14, Appl
26	121	2.3	1315	US-09-990-046-10	Sequence 10, Appl
27	117.5	2.2	2045	US-09-736-968A-109	Sequence 109, App
28	117.5	2.2	2045	US-09-736-969A-95	Sequence 95, Appl
29	117.5	2.2	2045	US-09-736-960-92	Sequence 92, Appl
30	116.5	2.2	3798	US-10-014-717-6	Sequence 6, Appl
31	116	2.2	1147	US-10-259-864-1	Sequence 1, Appl
32	115	2.2	1012	US-09-876-527-16	Sequence 16, Appl
33	114	2.2	979	US-10-038-686-3	Sequence 3, Appl
34	114	2.2	1023	US-09-893-519A-14	Sequence 14, Appl
35	111.5	2.1	1698	US-10-080-943-4	Sequence 4, Appl
36	111	2.1	3056	US-08-984-090-2	Sequence 2, Appl
37	111	2.1	3056	US-10-307-077-1	Sequence 1, Appl
38	111	2.1	3056	US-10-059-585-41	Sequence 41, Appl
39	111	2.1	3056	US-10-175-225-2	Sequence 2, Appl
40	110.5	2.1	621	US-09-799-777-77	Sequence 7, Appl
41	110.5	2.1	3057	US-10-175-225-3	Sequence 3, Appl
42	108.5	2.1	633	US-09-815-242-12077	Sequence 12077, A
43	108.5	2.1	1189	US-09-738-626-4140	Sequence 4140, Ap
44	108	2.1	852	US-10-261-482-2	Sequence 2, Appl
45	108	2.1	852	US-10-282-837-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-10-033-245-22
Sequence 22, Application US/10033245
Patent No. US20020160392A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P293081C7
CURRENT APPLICATION NUMBER: US/10/033,245
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965

;; PRIOR FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: 60/162,506
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: 60/170,262
;; PRIOR FILING DATE: 1999-12-09
;; PRIOR APPLICATION NUMBER: 60/187,202
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: 1999-06-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28634
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/28551
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: 2000-02-11
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: 2000-02-22
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: 2000-03-30
;; PRIOR APPLICATION NUMBER: PCT/US00/14941
;; PRIOR FILING DATE: 2000-05-30
;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; NUMBER OF SEQ ID NOS: 38
;; SEQ ID NO 22
;; LENGTH: 1029
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-033-245-22

Query Match 100.0%; Score 5241; DB 9; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHLVYVAMVILTLGPRADDSEFQALDIWPEEKPTAFIVDTSEALLPDWLKL 60
DB 1 MHLVYVAMVILTLGPRADDSEFQALDIWPEEKPTAFIVDTSEALLPDWLKL 60
QY 61 RMIRSEVLRVDAALDLEQOQLLFVQSGIPVSSMSKLLQFLDQAVADHPOTLEONIM 120
DB 61 RMIRSEVLRVDAALDLEQOQLLFVQSGIPVSSMSKLLQFLDQAVADHPOTLEONIM 120
QY 121 DKNYMAHLVEVQHERGASGGQTFHSLTASLPPRRDSTEAPKPSPEOPIGQIRIVGT 180
DB 121 DKNYMAHLVEVQHERGASGGQTFHSLTASLPPRRDSTEAPKPSPEOPIGQIRIVGT 180
QY 181 QLRVLGPEDDLAMFLQIFPLSPDPRMOWSSPPRVALALQALGOELARVVOGSPVPGI 240
DB 181 QLRVLGPEDDLAMFLQIFPLSPDPRMOWSSPPRVALALQALGOELARVVOGSPVPGI 240
QY 241 TVRVYLAALATLSSPHGALVMSHRSFPLACPLLRLOLCYORCPVDGTGESSLFLKVL 300
DB 241 TVRVYLAALATLSSPHGALVMSHRSFPLACPLLRLOLCYORCPVDGTGESSLFLKVL 300
QY 301 QMLQWLDSPEVEGGPLRAOLRLASQASAGRRLSDVVGGLRLAEALAFRODLEVVSTV 360
DB 301 QMLQWLDSPEVEGGPLRAOLRLASQASAGRRLSDVVGGLRLAEALAFRODLEVVSTV 360
QY 361 RAVIATIRSEOCSEVEDLISKVGLIEVRSPHLEELTAFTFSATADAASPPACKPVV 420
DB 361 RAVIATIRSEOCSEVEDLISKVGLIEVRSPHLEELTAFTFSATADAASPPACKPVV 420
QY 421 VVSSLIQEEPEPLAGKRGADGSLFAVRGLSPSSGILLVDMLEMDPEVSSCPDQIRLL 480
DB 421 VVSSLIQEEPEPLAGKRGADGSLFAVRGLSPSSGILLVDMLEMDPEVSSCPDQIRLL 480
QY 481 FSRKRGKGAQVPSFRYLLTLFTHOSSWPTLHQCIRVLLGKSRQGFDDBSASIDFLMAC 540
DB 481 FSRKRGKGAQVPSFRYLLTLFTHOSSWPTLHQCIRVLLGKSRQGFDDBSASIDFLMAC 540

QY 541 IHVPRIMQGRDQRTQKRRBELVLRVQGPBELISVLELLAEATRSODGDTAACSLIQAR 600
DB 541 IHVPRIMQGRDQRTQKRRBELVLRVQGPBELISVLELLAEATRSODGDTAACSLIQAR 600
QY 601 LPLLSCCGDDESVKRVTEHLSGCIQWGDVYGRRCRDLQLLYQRELPVPEVL 660
DB 601 LPLLSCCGDDESVKRVTEHLSGCIQWGDVYGRRCRDLQLLYQRELPVPEVL 660
QY 661 LHSEGAASSVCKLDGILHRTITLADTSDSRALENGADASMCRKLAVAHPPLLRHL 720
DB 661 LHSEGAASSVCKLDGILHRTITLADTSDSRALENGADASMCRKLAVAHPPLLRHL 720
QY 721 PMIAALLHGRTHLNFQEFROQNLSCFLVHYGLLELLQPHVFRSEHOGALWDCILSIRL 780
DB 721 PMIAALLHGRTHLNFQEFROQNLSCFLVHYGLLELLQPHVFRSEHOGALWDCILSIRL 780
QY 781 LFNVRSSRHIAAFINKEVQFIHKYTYNAPPAISFLQKHADPHDLSFQNSDLVMLKSL 840
DB 781 LFNVRSSRHIAAFINKEVQFIHKYTYNAPPAISFLQKHADPHDLSFQNSDLVMLKSL 840
QY 841 LAGLSLRDRTDRGDEEGEESAGSLPLVSVSLFTPLTAEMA PYMKRLSRGQTV 900
DB 841 LAGLSLRDRTDRGDEEGEESAGSLPLVSVSLFTPLTAEMA PYMKRLSRGQTV 900
QY 901 DLEVLSDIDEMSRRELLISFSTNQLRLMSSAECCRNLAFLSRMONSSTIAAF 960
DB 901 DLEVLSDIDEMSRRELLISFSTNQLRLMSSAECCRNLAFLSRMONSSTIAAF 960
QY 961 LPTFMYCLGSODFEVQVOTALRNLPEYALLCOEHAVALHRAFLVGMYGOMDPSQISEAL 1020
DB 961 LPTFMYCLGSODFEVQVOTALRNLPEYALLCOEHAVALHRAFLVGMYGOMDPSQISEAL 1020
QY 1021 RILHMEAVM 1029
DB 1021 RILHMEAVM 1029

RESULT 2
US-10-033-223-22
;; Sequence 22, Application US/10033223
;; Patent No. US20020164646A1
;; GENERAL INFORMATION:
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Ferreira, Napoleone
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: P2930R1C9
;; CURRENT APPLICATION NUMBER: US/10/033,223
;; CURRENT FILING DATE: 2001-12-27
;; PRIOR APPLICATION NUMBER: 60/095,325
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/112,851
;; PRIOR FILING DATE: 1998-12-16
;; PRIOR APPLICATION NUMBER: 60/113,145
;; PRIOR FILING DATE: 1998-12-16
;; PRIOR APPLICATION NUMBER: 60/113,511
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/115,558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115,565
;; PRIOR FILING DATE: 1999-01-12

```

: PRIOR APPLICATION NUMBER: 60/115,733
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/119,341
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: 60/119,537
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: 60/119,965
: PRIOR FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: 60/162,506
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: 60/170,262
: PRIOR FILING DATE: 1999-12-09
: PRIOR APPLICATION NUMBER: 60/187,202
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: PCT/US99/28634
: PRIOR FILING DATE: 1999-12-01
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US00/03565
: PRIOR FILING DATE: 2000-02-11
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: 2000-03-02
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: PCT/US00/14941
: PRIOR FILING DATE: 2000-05-30
: PRIOR APPLICATION NUMBER: PCT/US00/15264
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: 2000-12-01
: NUMBER OF SEQ ID NOS: 38
: SEQ ID NO 22
: LENGTH: 1029
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-033-223-22

```

```

Query Match      100.0%; Score 5241; DB 9; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHILVHAMVILITGPRADSEFOALDIMEPEEKPLPAFLVDTSEALLPDMKTL 60
DB 1 MHILVHAMVILITGPRADSEFOALDIMEPEEKPLPAFLVDTSEALLPDMKTL 60
QY 61 MIRSEVRLVDAALQDLEPQOLLFFVOSFGIPVSSMSKILQFLDQAVAHDPOTLEONIM 120
DB 61 MIRSEVRLVDAALQDLEPQOLLFFVOSFGIPVSSMSKILQFLDQAVAHDPOTLEONIM 120
QY 121 DKNTMAHLEVQHERGASGGOTFHSLLTASLPRRDSTEAPKPKSSPEOPIGGIRVGT 180
DB 121 DKNTMAHLEVQHERGASGGOTFHSLLTASLPRRDSTEAPKPKSSPEOPIGGIRVGT 180
QY 121 DKNTMAHLEVQHERGASGGOTFHSLLTASLPRRDSTEAPKPKSSPEOPIGGIRVGT 180
DB 121 DKNTMAHLEVQHERGASGGOTFHSLLTASLPRRDSTEAPKPKSSPEOPIGGIRVGT 180
QY 181 QLRVIGPREDLAGMFLQIFPLSPDRMOSSSPRYALALQOLGELARVVOGSEPEVGI 240
DB 181 QLRVIGPREDLAGMFLQIFPLSPDRMOSSSPRYALALQOLGELARVVOGSEPEVGI 240
QY 241 TVRVLOALATLLSPHGALVMSHSHFLACPLLROLCQYRCVPODTGFSSFLKVL 300
DB 241 TVRVLOALATLLSPHGALVMSHSHFLACPLLROLCQYRCVPODTGFSSFLKVL 300
QY 301 QMLQWLDSPGVGGPRLAQLRLMLASQASAGRRLSDVREGLLRLAALAFRODLEVVSTV 360
DB 301 QMLQWLDSPGVGGPRLAQLRLMLASQASAGRRLSDVREGLLRLAALAFRODLEVVSTV 360
QY 361 RAVATATSSGOCSEVPLISKVLQGLIEVRSPHLEELLTAFFSTAAASFPACKPV 420
DB 361 RAVATATSSGOCSEVPLISKVLQGLIEVRSPHLEELLTAFFSTAAASFPACKPV 420

```

```

QY 421 VVSSLLQEEBPLAGKKGADGSLAVALRGSSGLVDWLEMLDPEVSSCPDQLRL 480
DB 421 VVSSLLQEEBPLAGKKGADGSLAVALRGSSGLVDWLEMLDPEVSSCPDQLRL 480
QY 481 FSRKRGQAQVPSFRPYLLLETHQSSWPTLHQCIRVLLGKSRQRPDSASLDPLMAC 540
DB 481 FSRKRGQAQVPSFRPYLLLETHQSSWPTLHQCIRVLLGKSRQRPDSASLDPLMAC 540
QY 541 IHVPRINGROQRTQKRRRELYLVVQGPBELISVLELLIAEFETSSOGDPAACSLIAR 600
DB 541 IHVPRINGROQRTQKRRRELYLVVQGPBELISVLELLIAEFETSSOGDPAACSLIAR 600
QY 541 IHVPRINGROQRTQKRRRELYLVVQGPBELISVLELLIAEFETSSOGDPAACSLIAR 600
DB 541 IHVPRINGROQRTQKRRRELYLVVQGPBELISVLELLIAEFETSSOGDPAACSLIAR 600
QY 601 LPLLSCCGDESVRYKTEHLSCIOQMGDSVLGRRCRDLLQLYLQRPETRVPEVL 660
DB 601 LPLLSCCGDESVRYKTEHLSCIOQMGDSVLGRRCRDLLQLYLQRPETRVPEVL 660
QY 661 LHSEGAASSVCKIDGLIHRFTLLADTSDSRALENRGADASMAGRKLAVAHPLLRL 720
DB 661 LHSEGAASSVCKIDGLIHRFTLLADTSDSRALENRGADASMAGRKLAVAHPLLRL 720
QY 721 PMIAALIGRTHLNFQFRQONHLSCLFVGLLELQPHVFRSHOGALMDCILSFIRL 780
DB 721 PMIAALIGRTHLNFQFRQONHLSCLFVGLLELQPHVFRSHOGALMDCILSFIRL 780
QY 781 LLNRYKSSRHLLAFLNKKVQFTHKITYTNAAPAIISFLQKHADPLDLSEFNSDLYMLKSL 840
DB 781 LLNRYKSSRHLLAFLNKKVQFTHKITYTNAAPAIISFLQKHADPLDLSEFNSDLYMLKSL 840
QY 841 LAGLSLPSRDRDTRDGLDEGEESAGSLPLVSVSLFTPLTAEMAPYMKRLSGQYVE 900
DB 841 LAGLSLPSRDRDTRDGLDEGEESAGSLPLVSVSLFTPLTAEMAPYMKRLSGQYVE 900
QY 901 DLLEVLSDIDEMSRRRPETSFEFTNLQRLMSAEECCRNLAFLSIALSNQSPSIAAF 960
DB 901 DLLEVLSDIDEMSRRRPETSFEFTNLQRLMSAEECCRNLAFLSIALSNQSPSIAAF 960
QY 961 LPTFMYCIGSODFEVYQALRLNPEYALICQEAHVALHRAFLVGMGMDPSAQISEAL 1020
DB 961 LPTFMYCIGSODFEVYQALRLNPEYALICQEAHVALHRAFLVGMGMDPSAQISEAL 1020
QY 1021 RILHMEAVM 1029
DB 1021 RILHMEAVM 1029

```

RESULT 3
US-10-033-167-22

```

: Sequence 22, Application US/10033167
: Publication No. US20020182618A1
: GENERAL INFORMATION:
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2930R1C10
: CURRENT FILING DATE: 2001-12-27
: PRIOR APPLICATION NUMBER: US/10/033,167
: PRIOR FILING DATE: 2001-12-27
: PRIOR APPLICATION NUMBER: 60/095,325
: PRIOR FILING DATE: 1998-08-04
: PRIOR APPLICATION NUMBER: 60/112,851
: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,145

```

```

PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 22
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-10-033-167-22

Query Match      100.0%; Score 5241; DB 9; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHILVNAVWILTLGPRADSEFOALDLIWPBEKPLPTAFVDTSEELLPDMLKL 60
    |||||
DB 1 MHILVNAVWILTLGPRADSEFOALDLIWPBEKPLPTAFVDTSEELLPDMLKL 60
    |||||
QY 61 MTRSEVLRVDAALDLEPQQLLFVQSGEIPVSSMSKLIQFLDOAVAHDPQLEQNM 120
    |||||
DB 61 MTRSEVLRVDAALDLEPQQLLFVQSGEIPVSSMSKLIQFLDOAVAHDPQLEQNM 120
    |||||
QY 121 DKNMALLVEONERKSGSGGOTFHSLLTASIPPRDSTEAPKPSSEOPRIGORIRVGT 180
    |||||
DB 121 DKNMALLVEONERKSGSGGOTFHSLLTASIPPRDSTEAPKPSSEOPRIGORIRVGT 180
    |||||
QY 181 QLRLVGEDDLAGMFLQIFPLSPDRMSSPREVALALQALGOELARVVOGSPVPGI 240
    |||||
DB 181 QLRLVGEDDLAGMFLQIFPLSPDRMSSPREVALALQALGOELARVVOGSPVPGI 240
    |||||
QY 241 TVRYLQALATLLSSPHGALVMSHNSHFLACPLRLQOLCOYQRCVPQDTGFSFLKVL 300
    |||||
DB 241 TVRYLQALATLLSSPHGALVMSHNSHFLACPLRLQOLCOYQRCVPQDTGFSFLKVL 300
    |||||
QY 301 QMLOWDSPEVEGRLRAQLRLMLASQASAGRRLSDVNGGLRLALAFRQDELVSSIV 360
    |||||
```

```

|||||
DB 301 QMLOWDSPEVEGRLRAQLRLMLASQASAGRRLSDVNGGLRLALAFRQDELVSSIV 360
    |||||
QY 361 RAVATLRSGQCSVEPDLISKVLOGILEVRSPIHEELTLTAFSATADAASPACPVV 420
    |||||
DB 361 RAVATLRSGQCSVEPDLISKVLOGILEVRSPIHEELTLTAFSATADAASPACPVV 420
    |||||
QY 421 VVSSLLOEEEPBLAGKRGADGSLLEAVRLGPSSGGLLVDMLEMDPEVSSCPDLQLRL 480
    |||||
DB 421 VVSSLLOEEEPBLAGKRGADGSLLEAVRLGPSSGGLLVDMLEMDPEVSSCPDLQLRL 480
    |||||
QY 481 FSRKRGQAVPSFRPLTLFTHOSWPTLHOCIVLLGKSRQRPDPASLDPLMAC 540
    |||||
DB 481 FSRKRGQAVPSFRPLTLFTHOSWPTLHOCIVLLGKSRQRPDPASLDPLMAC 540
    |||||
QY 541 IHVRIMQGRQRTPOKRREELVRYVGPPELISVVELLAEAFERSODGDTAACSLQAR 600
    |||||
DB 541 IHVRIMQGRQRTPOKRREELVRYVGPPELISVVELLAEAFERSODGDTAACSLQAR 600
    |||||
QY 601 LPLLLSCCGDDESVRKYTEHLSCGICQWGDVIGRRCDLLQLYLQRPFLRVPEVL 660
    |||||
DB 601 LPLLLSCCGDDESVRKYTEHLSCGICQWGDVIGRRCDLLQLYLQRPFLRVPEVL 660
    |||||
QY 661 LHSEGAASSVCKLDGLIHRTITLLADTSDSRALENGADASMACKRLAVAHPLLRHL 720
    |||||
DB 661 LHSEGAASSVCKLDGLIHRTITLLADTSDSRALENGADASMACKRLAVAHPLLRHL 720
    |||||
QY 721 PMIAALLHGRTHLNFQERQONHLSCEFLHYGLLELLQPHYFSEHOGALWDCLSFTRL 780
    |||||
DB 721 PMIAALLHGRTHLNFQERQONHLSCEFLHYGLLELLQPHYFSEHOGALWDCLSFTRL 780
    |||||
QY 781 LLNTRKSSRHIAFTINKFVOEIHKITYYNAPASISFLQKADPHLDSFDNSDLVMLKSL 840
    |||||
DB 781 LLNTRKSSRHIAFTINKFVOEIHKITYYNAPASISFLQKADPHLDSFDNSDLVMLKSL 840
    |||||
QY 841 LAGLSRSDRDRTRGIDEESEESSAGSLPVSLSFTPLTAEMAPYMKRLSRGQTV 900
    |||||
DB 841 LAGLSRSDRDRTRGIDEESEESSAGSLPVSLSFTPLTAEMAPYMKRLSRGQTV 900
    |||||
QY 901 DLLEVLSDIDEMSRRELLISFFSTNLQRLMSAECCRNIAFSLATRSKMSNPSIAAF 960
    |||||
DB 901 DLLEVLSDIDEMSRRELLISFFSTNLQRLMSAECCRNIAFSLATRSKMSNPSIAAF 960
    |||||
QY 961 LPTMYCLGSDPEFVYOTALNLPYVALLCOENAAVLHRAFLVGMVGOMDPSAOISEAL 1020
    |||||
DB 961 LPTMYCLGSDPEFVYOTALNLPYVALLCOENAAVLHRAFLVGMVGOMDPSAOISEAL 1020
    |||||
QY 1021 RILHMEAVM 1029
    |||||
DB 1021 RILHMEAVM 1029
    |||||

RESULT 4
US-10-033-244-22
Sequence 22, Application US/10033244
Publication No. US2002019266A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1C2
```

```

; CURRENT APPLICATION NUMBER: US/10/033,244
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 22
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-244-22

```

```

Query Match      100.0%; Score 5241; DB 9; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MHLVYHAWYILTLGPRRDSDFQALDDIMFPEEKPLPTATLVDTSEALLPWLKL 60
DB 1 MHLVYHAWYILTLGPRRDSDFQALDDIMFPEEKPLPTATLVDTSEALLPWLKL 60
QY 61 RMIRSEVLRLVDAALODLEPQQLLFVQSFQIVVSSMSKILQPLDOVAHDPOTLEONIM 120
DB 61 RMIRSEVLRLVDAALODLEPQQLLFVQSFQIVVSSMSKILQPLDOVAHDPOTLEONIM 120
QY 121 DKNYMAHVEVOHERGASGGQTFHSLTASLPERRDSTEAPKPKSSPEODIGGRIVGT 180
DB 121 DKNYMAHVEVOHERGASGGQTFHSLTASLPERRDSTEAPKPKSSPEODIGGRIVGT 180
QY 181 QLRVLGPEDDLAGMFIQITPLSPDPRWSSSPPVALLQALGOELARVQSSPEVPGI 240
DB 181 QLRVLGPEDDLAGMFIQITPLSPDPRWSSSPPVALLQALGOELARVQSSPEVPGI 240

```

```

DB 181 QLRVLGPEDDLAGMFIQITPLSPDPRWSSSPPVALLQALGOELARVQSSPEVPGI 240
QY 241 TVEVLQALATLLSPHGALVMSHSHFLACPLRLQCYORCVODTFSSLEFLKVL 300
DB 241 TVEVLQALATLLSPHGALVMSHSHFLACPLRLQCYORCVODTFSSLEFLKVL 300
QY 301 QMLQWLDSPGVEGGLPLRAQLRMLASQASAGRRLSDVNGGLRLAEALAFRODEEVSSIV 360
DB 301 QMLQWLDSPGVEGGLPLRAQLRMLASQASAGRRLSDVNGGLRLAEALAFRODEEVSSIV 360
QY 361 RAVIATLRSGEQSVPEPDLISKVLQGLIEVRSHPLEELLTAFFSATAADAAPPACKPVY 420
DB 361 RAVIATLRSGEQSVPEPDLISKVLQGLIEVRSHPLEELLTAFFSATAADAAPPACKPVY 420
QY 421 VVSSLLQEEEPPLAGGKPGADGSLAEAVRGPSSGLVMDLEMDPEVSSCCDLQRL 480
DB 421 VVSSLLQEEEPPLAGGKPGADGSLAEAVRGPSSGLVMDLEMDPEVSSCCDLQRL 480
QY 481 FSRKKGQAQVPSFRPYLLTFTHOSWPTLHQCIKRVLLGKSGREDFPSASLDFLMAC 540
DB 481 FSRKKGQAQVPSFRPYLLTFTHOSWPTLHQCIKRVLLGKSGREDFPSASLDFLMAC 540
QY 541 IHVPRIWQGDQRTPOKRRBELVLRVQGPPELLISVELLTAEAETRSODGTACSLIQAR 600
DB 541 IHVPRIWQGDQRTPOKRRBELVLRVQGPPELLISVELLTAEAETRSODGTACSLIQAR 600
QY 601 LPPLSCCGDDSVKRVTEHLSGCIQWGDVYLRGRCDLLQLYLQRELPVPEVL 660
DB 601 LPPLSCCGDDSVKRVTEHLSGCIQWGDVYLRGRCDLLQLYLQRELPVPEVL 660
QY 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRLAVAHPLLLRL 720
DB 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRLAVAHPLLLRL 720
QY 721 PMTALLHGTHLNFQEFROHNLSCFLHYLGLELLQHPVFSSEHGALMDCLSLTRL 780
DB 721 PMTALLHGTHLNFQEFROHNLSCFLHYLGLELLQHPVFSSEHGALMDCLSLTRL 780
QY 781 LKNYRKSSRLAAFINKEVQFIHKYIYVNAAPAAISFLQKHADPLHDSPNSDLVMLKSL 840
DB 781 LKNYRKSSRLAAFINKEVQFIHKYIYVNAAPAAISFLQKHADPLHDSPNSDLVMLKSL 840
QY 841 LAGLSLPSRDDRTDGLDEGEBSAGSLPLYVSLFTPLTAEMAAPYKRLSRGCTVE 900
DB 841 LAGLSLPSRDDRTDGLDEGEBSAGSLPLYVSLFTPLTAEMAAPYKRLSRGCTVE 900
QY 901 DLLEVSDIDEMRRREILISFSTNLORLMSSAECCRNLAFLSRMONSPSTAAAF 960
DB 901 DLLEVSDIDEMRRREILISFSTNLORLMSSAECCRNLAFLSRMONSPSTAAAF 960
QY 961 LPTFMVCLGSQDFEVVOTALRNLPEVALLCOEHAVALLHRAFLVGMYGMDPSNOISEAL 1020
DB 961 LPTFMVCLGSQDFEVVOTALRNLPEVALLCOEHAVALLHRAFLVGMYGMDPSNOISEAL 1020
QY 1021 RILHMEAVM 1029
DB 1021 RILHMEAVM 1029

```

```

RESULT 5
US-10-033-435-22
; Sequence 22, Application US/10033435
; Publication No. US20030027256A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann

```

```

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1C5
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 22
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-10-033-435-22
Query Match 100.0%; Score 5241; DB 9; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 121 DKNYMAHLVEYQHERGASGGOTFHSLLTASLPERRDSTEAPKPSSEOPIGCRIRVGT 180
DB 121 DKNYMAHLVEYQHERGASGGOTFHSLLTASLPERRDSTEAPKPSSEOPIGCRIRVGT 180
QY 181 QLRVAGEPDDLACMFLOIFPLSPDPRWMOSSPRVALALQOALQOELARVYOGSPVPYGI 240
DB 181 QLRVAGEPDDLACMFLOIFPLSPDPRWMOSSPRVALALQOALQOELARVYOGSPVPYGI 240
QY 241 TVRVLOALATLSSPHGALVMSHRSHPFLACPLRLQCOVORCPVDDTGFSSFLKVL 300
DB 241 TVRVLOALATLSSPHGALVMSHRSHPFLACPLRLQCOVORCPVDDTGFSSFLKVL 300
QY 301 QMLQWLDSPGVGGPPLRAOQLMLASQASAGRRLSDVAGLLRLAEALAFRODLEEVSV 360
DB 301 QMLQWLDSPGVGGPPLRAOQLMLASQASAGRRLSDVAGLLRLAEALAFRODLEEVSV 360
QY 361 RAVIATLRSGEQCEVEDLSKVLQGLIEVRSPLLEELTAFFSATADAAFPACRPVY 420
DB 361 RAVIATLRSGEQCEVEDLSKVLQGLIEVRSPLLEELTAFFSATADAAFPACRPVY 420
QY 421 VVSSLIQEEEPFLAGKRGADGSLAVRGLSPSSGLVDMLEMDPEVSSCPDLQRL 480
DB 421 VVSSLIQEEEPFLAGKRGADGSLAVRGLSPSSGLVDMLEMDPEVSSCPDLQRL 480
QY 481 FSRKRGQAQAVPSFRPYLLTFTHQSSWPTLHQCIRVLLGKSRQRFDPASIDFLMAC 540
DB 481 FSRKRGQAQAVPSFRPYLLTFTHQSSWPTLHQCIRVLLGKSRQRFDPASIDFLMAC 540
QY 541 IHVRIMQGRQRTPOKRREELVLRVGPPELISIVELILEAEFRSDGDTAACSLQAR 600
DB 541 IHVRIMQGRQRTPOKRREELVLRVGPPELISIVELILEAEFRSDGDTAACSLQAR 600
QY 601 LPILLSCCCGDBESVRKYTEHLSGCIQWGDPSVLRGRCDLLOLYQREPLRVPEVL 660
DB 601 LPILLSCCCGDBESVRKYTEHLSGCIQWGDPSVLRGRCDLLOLYQREPLRVPEVL 660
QY 661 LHSEGAASSVCKLDGLIHRFTYLLADTSDSRALENRGADASMACRLAVALPILLRL 720
DB 661 LHSEGAASSVCKLDGLIHRFTYLLADTSDSRALENRGADASMACRLAVALPILLRL 720
QY 721 PMTALLHGRTHLNFQERROONHLSCLFHVYGLLELLOPHYFRSEHOGALMDCILSIFRL 780
DB 721 PMTALLHGRTHLNFQERROONHLSCLFHVYGLLELLOPHYFRSEHOGALMDCILSIFRL 780
QY 781 LLNTRKSSRHIAFINKFVOFIHKYITYNAPAAISFLQKADPLHDUSFNSDLVMLKSL 840
DB 781 LLNTRKSSRHIAFINKFVOFIHKYITYNAPAAISFLQKADPLHDUSFNSDLVMLKSL 840
QY 841 LAGLSLPSRDDRTDRCGLDEGEESAGSLPVVSFLPTLTAEMAPYKRLRSRGQTV 900
DB 841 LAGLSLPSRDDRTDRCGLDEGEESAGSLPVVSFLPTLTAEMAPYKRLRSRGQTV 900
QY 901 DLEVLSDIDEMSRREPELISFSTINQRLMSSECCRNIAFSLARSMONSSIAAF 960
DB 901 DLEVLSDIDEMSRREPELISFSTINQRLMSSECCRNIAFSLARSMONSSIAAF 960
QY 961 LPTFMVCLGSODEFEVOTALRNLPYVALLCOEHAVALHRAFLVGMYGMDPSAQISEAL 1020
DB 961 LPTFMVCLGSODEFEVOTALRNLPYVALLCOEHAVALHRAFLVGMYGMDPSAQISEAL 1020
QY 1021 RILHMEAVM 1029
DB 1021 RILHMEAVM 1029

```

```

RESULT 6
US-10-032-990-22
Sequence 22, Application US/10032990
Publication No. US20030032060A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc

```


APPLICANT: Ferrara, Napoleone
APPLICANT: Gao, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Collin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1C11
CURRENT APPLICATION NUMBER: US/10/032,990
PRIOR FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 22
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-10-032-990-22

Query Match 100.0% Score 5241: DB 9: Length 1029:
Best Local Similarity 100.0% Pred. No. 0:
Matches 1029: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 MHILVHNAVILITLGPRAADSEFOALLDIWFPEEKPLPTAFVYDTSEALLPDWLK 60
DB 1 MHILVHNAVILITLGPRAADSEFOALLDIWFPEEKPLPTAFVYDTSEALLPDWLK 60
QY 61 RMIRSEVRLVDAIODELPQOLLLFVOSFGIPVSSMSKLPLODAVAHDQTEQNIM 120
DB 61 RMIRSEVRLVDAIODELPQOLLLFVOSFGIPVSSMSKLPLODAVAHDQTEQNIM 120
QY 121 DKNYMAHVEVOHEGASGQTFHSLTASLPRRRSTAPKPKSSPEQITQGRIRVGT 180
DB 121 DKNYMAHVEVOHEGASGQTFHSLTASLPRRRSTAPKPKSSPEQITQGRIRVGT 180
QY 181 QLRVLGPEDDIAGMFLQIPIPLSPDPWQSSPRVALAQAOLGQELARVAGSEVPST 240
DB 181 QLRVLGPEDDIAGMFLQIPIPLSPDPWQSSPRVALAQAOLGQELARVAGSEVPST 240
QY 241 TVRVLOALATLLSPHGALVMSMRSHFLACPLRLQOLCOYORCVQPDGFSLSFLKVL 300
DB 241 TVRVLOALATLLSPHGALVMSMRSHFLACPLRLQOLCOYORCVQPDGFSLSFLKVL 300
QY 301 QMLQWIDSPGVGGPLRAQLRMLASQASGRUSDVGGILRLAEALARODLEVSSIV 360
DB 301 QMLQWIDSPGVGGPLRAQLRMLASQASGRUSDVGGILRLAEALARODLEVSSIV 360
QY 361 RAVIATLRSGECSVEPDLISKVLOGLIEVRSBHLELLTAFFSATADAASFPPACKPVY 420
DB 361 RAVIATLRSGECSVEPDLISKVLOGLIEVRSBHLELLTAFFSATADAASFPPACKPVY 420
QY 421 VVSSLLLOEEBPLAGKPPADGGSLEAVRLGSSGGLVVMLEMLDEPVVSSCPDLORL 480
DB 421 VVSSLLLOEEBPLAGKPPADGGSLEAVRLGSSGGLVVMLEMLDEPVVSSCPDLORL 480
QY 481 FSRKRGGAQVPSFRPYLLTFTHOSSMPTLHOCIRVLLGSRREPRPSADLTLMAC 540
DB 481 FSRKRGGAQVPSFRPYLLTFTHOSSMPTLHOCIRVLLGSRREPRPSADLTLMAC 540
QY 541 HVPRLMOGRDQTPQKRREELVLRVQPELISVLELLAEATRSQDQDTAACSILQAR 600
DB 541 HVPRLMOGRDQTPQKRREELVLRVQPELISVLELLAEATRSQDQDTAACSILQAR 600
QY 601 LPILLSCCGDDESVAKVTEHLSGCIQWGDVYLGRCDLLQLLYQRPRLVPPEVL 660
DB 601 LPILLSCCGDDESVAKVTEHLSGCIQWGDVYLGRCDLLQLLYQRPRLVPPEVL 660
QY 661 LHSSEGAASSVCKLQGLIHRFTLLADTSDSRALENRGADASACRKLVAHPDLLRLRL 720
DB 661 LHSSEGAASSVCKLQGLIHRFTLLADTSDSRALENRGADASACRKLVAHPDLLRLRL 720
QY 721 PMIAALLHGRTHLNFQEFROQNLSCFLHVLGILLELQHPVRSERHOGALMDCILSFIRL 780
DB 721 PMIAALLHGRTHLNFQEFROQNLSCFLHVLGILLELQHPVRSERHOGALMDCILSFIRL 780
QY 781 LLYNRKSSRHAAFLINKFVQFIHKYTYVNAPAALSTLQKHADPLHDLSFSDNSDVLWKL 840
DB 781 LLYNRKSSRHAAFLINKFVQFIHKYTYVNAPAALSTLQKHADPLHDLSFSDNSDVLWKL 840
QY 841 LAGLSLPSRDDPTDRGLDEGEESAGSLPIYSVSLFPLTAEMAAPMKRLSRQCTYE 900
DB 841 LAGLSLPSRDDPTDRGLDEGEESAGSLPIYSVSLFPLTAEMAAPMKRLSRQCTYE 900
QY 901 DLLEVLSDIDEMSRRRPELISFEFTNLQRLMSAECCCNLAFSLARSMQNSPSTAAAF 960
DB 901 DLLEVLSDIDEMSRRRPELISFEFTNLQRLMSAECCCNLAFSLARSMQNSPSTAAAF 960
QY 961 LPTFMTCGSDQFEVVQTLRLNLPYALLCOEHAVALHRAFLVGYGMQDPSAQISEAL 1020
DB 961 LPTFMTCGSDQFEVVQTLRLNLPYALLCOEHAVALHRAFLVGYGMQDPSAQISEAL 1020
QY 1021 RILHMEAVM 1029
DB 1021 RILHMEAVM 1029

```

RESULT 7
US-09-929-769-7
: Sequence 7, Application US/09929769
: Publication No. US20030055224A1
: GENERAL INFORMATION:
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Polakis, Paul
: APPLICANT: Shou, Jianyong
: APPLICANT: Smith, Victoria
: APPLICANT: Soriano, Robert
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wu, Thomas D.
: APPLICANT: Zhang, Zhenli
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
: FILE REFERENCE: P5007R1-US
: CURRENT APPLICATION NUMBER: US/09/929,769
: CURRENT FILING DATE: 2001-08-14
: PRIOR APPLICATION NUMBER: 60/089,653
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/090,355
: PRIOR FILING DATE: 1998-06-23
: PRIOR APPLICATION NUMBER: 60/104,257
: PRIOR FILING DATE: 1998-10-14
: PRIOR APPLICATION NUMBER: 60/119,537
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: 60/141,037
: PRIOR FILING DATE: 1999-06-23
: PRIOR APPLICATION NUMBER: 60/145,698
: PRIOR FILING DATE: 1999-07-26
: PRIOR APPLICATION NUMBER: 60/162,506
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: PCT/US99/20111
: PRIOR FILING DATE: 1999-09-01
: PRIOR APPLICATION NUMBER: PCT/US99/28634
: PRIOR FILING DATE: 1999-12-01
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US00/00219
: PRIOR FILING DATE: 2000-01-05
: PRIOR APPLICATION NUMBER: PCT/US00/00376
: PRIOR FILING DATE: 2000-01-06
: PRIOR APPLICATION NUMBER: PCT/US00/04342
: PRIOR FILING DATE: 2000-02-18
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: PCT/US00/13705
: PRIOR FILING DATE: 2000-05-17
: PRIOR APPLICATION NUMBER: PCT/US00/23328
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: PCT/US01/06520
: PRIOR FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: PCT/US01/20118
: PRIOR FILING DATE: 2001-06-22
: PRIOR APPLICATION NUMBER: 09/888,257
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 8
: SEQ ID NO 7
: LENGTH: 1029
: TYPE: PRT
: ORGANISM: Homo Sapien
US-09-929-769-7

```

```

Query Match      100.0%: Score 5241: DB 9: Length 1029:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 1029: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

```

```

QY 1 MHILVHANYILTLGPRADSEFOALLDIWPEEKPLPTAFVDTSEALLLPDKL 60
DB 1 MHILVHANYILTLGPRADSEFOALLDIWPEEKPLPTAFVDTSEALLLPDKL 60

```

```

QY 61 RMIRSEVLRLVDAALDLEPOQLLEFVOSFCIPVSSMSKLLQFLDQAVAHDPOTLEONIM 120
DB 61 RMIRSEVLRLVDAALDLEPOQLLEFVOSFCIPVSSMSKLLQFLDQAVAHDPOTLEONIM 120
QY 121 DKNYMAHVEVQHERGASGGOTFHSLLTASLPRRROSTEARPKRKSSEQDTGGGRIRVGT 180
DB 121 DKNYMAHVEVQHERGASGGOTFHSLLTASLPRRROSTEARPKRKSSEQDTGGGRIRVGT 180
QY 181 QLRVLPGEEDDLACMFQIIFPLSPDPKRWSSPPVALALQALGOELARVQSSPEVPGI 240
DB 181 QLRVLPGEEDDLACMFQIIFPLSPDPKRWSSPPVALALQALGOELARVQSSPEVPGI 240
QY 241 TVRYLQALATLSSPHGALVMSMHRSHFLACPLRLQCYORCVDPDTGSSFLFKVL 300
DB 241 TVRYLQALATLSSPHGALVMSMHRSHFLACPLRLQCYORCVDPDTGSSFLFKVL 300
QY 301 QMLQWLDSPGVGGPLPNAQRLMLASQASAGRRLSDVAGGLRLALAEALAFODLEVSSTV 360
DB 301 QMLQWLDSPGVGGPLPNAQRLMLASQASAGRRLSDVAGGLRLALAEALAFODLEVSSTV 360
QY 361 RAVIATLRSGECCSVEBDLISKVLOGLIEVRSPLLELLTAFFSATADAASPPACKPVV 420
DB 361 RAVIATLRSGECCSVEBDLISKVLOGLIEVRSPLLELLTAFFSATADAASPPACKPVV 420
QY 421 VVSSLIQEEEPPLAGKPGADGSLAVRIGPSSGLLVMDLMDPEVSSCPDLQRL 480
DB 421 VVSSLIQEEEPPLAGKPGADGSLAVRIGPSSGLLVMDLMDPEVSSCPDLQRL 480
QY 481 FSRKKGQAQVPSFRYLLTLFTHOSSWPTLHOCIVLLGKSGREGFPDSSASIDFLMAC 540
DB 481 FSRKKGQAQVPSFRYLLTLFTHOSSWPTLHOCIVLLGKSGREGFPDSSASIDFLMAC 540
QY 541 IHVPRIMQGRDQRTPOKRREELVLRVGPBELISVELLAEATRSODGDTAACSLIQAR 600
DB 541 IHVPRIMQGRDQRTPOKRREELVLRVGPBELISVELLAEATRSODGDTAACSLIQAR 600
QY 541 IHVPRIMQGRDQRTPOKRREELVLRVGPBELISVELLAEATRSODGDTAACSLIQAR 600
DB 541 IHVPRIMQGRDQRTPOKRREELVLRVGPBELISVELLAEATRSODGDTAACSLIQAR 600
QY 601 LPLILSCCGDDESVRKVTETHLSGCIQOWGDSVYGRRCROLLQLYQRELRVPVEVL 660
DB 601 LPLILSCCGDDESVRKVTETHLSGCIQOWGDSVYGRRCROLLQLYQRELRVPVEVL 660
QY 661 LHSEGAASSVCKLDGLIHREFTLLADTSDSRALENGDASMACRKLAVANPLLRLL 720
DB 661 LHSEGAASSVCKLDGLIHREFTLLADTSDSRALENGDASMACRKLAVANPLLRLL 720
QY 721 PMIAALLHGRTHLNFQEFROQNLSCFLHVLGLLELLQPHVPNSEHOGALMDCILSIRL 780
DB 721 PMIAALLHGRTHLNFQEFROQNLSCFLHVLGLLELLQPHVPNSEHOGALMDCILSIRL 780
QY 781 LLYNRKSSRHLLAFINKEVQFIHKYITYNAPAIISFLQKHADPLHDLSPNSDLVMLKSL 840
DB 781 LLYNRKSSRHLLAFINKEVQFIHKYITYNAPAIISFLQKHADPLHDLSPNSDLVMLKSL 840
QY 841 LAGLSLPSRDDRDGDEEGEESASGLPIVSVSLFTPLTAEMAPYMKRRLSGQTV 900
DB 841 LAGLSLPSRDDRDGDEEGEESASGLPIVSVSLFTPLTAEMAPYMKRRLSGQTV 900
QY 901 DLEVLSDIDEMSRREPILISFESTNQLRLMSAEECCRMILASLARSQNSPSTIAAF 960
DB 901 DLEVLSDIDEMSRREPILISFESTNQLRLMSAEECCRMILASLARSQNSPSTIAAF 960
QY 961 LPTFMYCLGSODEFVQVOTALRNLPEYALLCOEHAVALLHRAFLVGMVGOMDPSAQISEAL 1020
DB 961 LPTFMYCLGSODEFVQVOTALRNLPEYALLCOEHAVALLHRAFLVGMVGOMDPSAQISEAL 1020
QY 1021 RIILHMEAVM 1029
DB 1021 RIILHMEAVM 1029

```

```

RESULT 8
US-10-032-996-22
: Sequence 22, Application US/10032996
: Publication No. US20030054447A1

```

GENERAL INFORMATION:

APPLICANT: Botstein, David
 APPLICANT: Deshoyers, Luc
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Goddard, Audrey
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P293ORIC3
 CURRENT FILING DATE: 2001-12-27
 PRIOR APPLICATION NUMBER: 60/095,325
 PRIOR FILING DATE: 1998-08-04
 PRIOR APPLICATION NUMBER: 60/112,851
 PRIOR FILING DATE: 1998-12-16
 PRIOR APPLICATION NUMBER: 60/113,145
 PRIOR FILING DATE: 1998-12-16
 PRIOR APPLICATION NUMBER: 60/113,511
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/115,558
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115,565
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115,733
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/119,341
 PRIOR FILING DATE: 1999-02-09
 PRIOR APPLICATION NUMBER: 60/119,537
 PRIOR FILING DATE: 1999-02-10
 PRIOR APPLICATION NUMBER: 60/119,965
 PRIOR FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: 60/162,506
 PRIOR FILING DATE: 1999-10-29
 PRIOR APPLICATION NUMBER: 60/170,262
 PRIOR FILING DATE: 1999-12-09
 PRIOR APPLICATION NUMBER: 60/187,202
 PRIOR FILING DATE: 2000-03-03
 PRIOR APPLICATION NUMBER: PCT/US99/12252
 PRIOR FILING DATE: 1999-06-02
 PRIOR APPLICATION NUMBER: PCT/US99/28634
 PRIOR FILING DATE: 1999-12-01
 PRIOR APPLICATION NUMBER: PCT/US99/28551
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US00/03565
 PRIOR FILING DATE: 2000-02-11
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: PCT/US00/05841
 PRIOR FILING DATE: 2000-03-02
 PRIOR APPLICATION NUMBER: PCT/US00/08439
 PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: PCT/US00/14941
 PRIOR FILING DATE: 2000-05-30
 PRIOR APPLICATION NUMBER: PCT/US00/15264
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: 2000-12-01
 NUMBER OF SEQ ID NOS: 38
 SEQ ID NO 22
 LENGTH: 1029
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-032-996-22

Query Match

100.0%; Score 5241; DB 9; Length 1029;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MHLIVYHAWVILITTCGPRPADSEFPALLDIMPREKPLPTAFIVNTSEALLIPWKTL	60
DB	1	MHLVYHAWVILITTCGPRPADSEFPALLDIMPREKPLPTAFIVNTSEALLIPWKTL	60
QY	61	RMIRSEVLRYDAALODLEFPQDLLEFVQSGFIPVSSMSKLLFLDQAVAHDPQLEONIM	120
DB	61	RMIRSEVLRYDAALODLEFPQDLLEFVQSGFIPVSSMSKLLFLDQAVAHDPQLEONIM	120
QY	121	DKNYMAHVEVOYHERGASGGQTFHSLITLTPASLPERRDSTEAPRKSSPEQIGQGRIRVGT	180
DB	121	DKNYMAHVEVOYHERGASGGQTFHSLITLTPASLPERRDSTEAPRKSSPEQIGQGRIRVGT	180
QY	181	QLRVLPEDDLAGMFLQITPPLSPDPFWQSSSRPVAAIQALGQGLARVYQSGSPVPGI	240
DB	181	QLRVLPEDDLAGMFLQITPPLSPDPFWQSSSRPVAAIQALGQGLARVYQSGSPVPGI	240
QY	241	TVRVQALATLLSPHGGALVMSMHSRHLACPLRLQOLCOYRCVQDTGFSSFLFKVLL	300
DB	241	TVRVQALATLLSPHGGALVMSMHSRHLACPLRLQOLCOYRCVQDTGFSSFLFKVLL	300
QY	301	OMLQWLDSPEVSGGFLRAQLRLMLASQASAGRISDVRGGLRLAEALAEFRODLEVVSSTV	360
DB	301	OMLQWLDSPEVSGGFLRAQLRLMLASQASAGRISDVRGGLRLAEALAEFRODLEVVSSTV	360
QY	361	RAVIATLRSGECCSVPEPDLISKVLOGLIEVRSHLEELLTAFPSATADAAPPACKPVV	420
DB	361	RAVIATLRSGECCSVPEPDLISKVLOGLIEVRSHLEELLTAFPSATADAAPPACKPVV	420
QY	421	VVSSLLQEEEPPLAGSKPADGSLFAVRLGPSGLVVMLEMLDEEVVSSCGDLOLRLL	480
DB	421	VVSSLLQEEEPPLAGSKPADGSLFAVRLGPSGLVVMLEMLDEEVVSSCGDLOLRLL	480
QY	481	FSRRKKGQAOYVSPFRPYLLTFTHQSSWPTLHQICIRVLLGSRERFDPSSALDELMAC	540
DB	481	FSRRKKGQAOYVSPFRPYLLTFTHQSSWPTLHQICIRVLLGSRERFDPSSALDELMAC	540
QY	541	IHYPRIMQGRDQRTPOKREELVLRVQGPGLISVLELLIAEATRSQDDDTAACSILQAR	600
DB	541	IHYPRIMQGRDQRTPOKREELVLRVQGPGLISVLELLIAEATRSQDDDTAACSILQAR	600
QY	601	LPLLSCCGGDESVKRYVEHLSGCIQWGDVSLGRCRDLLQLYLRPELTPVPEVL	660
DB	601	LPLLSCCGGDESVKRYVEHLSGCIQWGDVSLGRCRDLLQLYLRPELTPVPEVL	660
QY	661	LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMAACKLAVAHPLLLRLH	720
DB	661	LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMAACKLAVAHPLLLRLH	720
QY	721	PMIAALLHGRTHINPOEFQOQNLSCFLVHLGLLELLQHPHVESEHGALMDCLLSFIRL	780
DB	721	PMIAALLHGRTHINPOEFQOQNLSCFLVHLGLLELLQHPHVESEHGALMDCLLSFIRL	780
QY	781	LNTRYKSRHLLAFAIFKPFQFIHKYTYNAPAAISFLQKHADPLHDLSDNSDLVLEKSL	840
DB	781	LNTRYKSRHLLAFAIFKPFQFIHKYTYNAPAAISFLQKHADPLHDLSDNSDLVLEKSL	840
QY	841	LAGLSLPSRDDRTDRGLDEEGEESAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQIVE	900
DB	841	LAGLSLPSRDDRTDRGLDEEGEESAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQIVE	900
QY	901	DLLEVLSDIDEMSRRRPELTSFSTNLRQMSAECCCNLAISLARQMONSPSTAAAF	960
DB	901	DLLEVLSDIDEMSRRRPELTSFSTNLRQMSAECCCNLAISLARQMONSPSTAAAF	960
QY	961	LPTFMVCLSGDPEFVQVQALRLNPEYALLCQEHAAVLLHRAFLVGYGQMDPSAQISEAL	1020
DB	961	LPTFMVCLSGDPEFVQVQALRLNPEYALLCQEHAAVLLHRAFLVGYGQMDPSAQISEAL	1020
QY	1021	RIIHMEAVM 1029	
DB	1021	RIIHMEAVM 1029	

Db 1021 RILHMEAVM 1029

RESULT 9
US-10-033-396-22

Sequence 22, Application US/10033396
 Publication No. US20030077657A1
 GENERAL INFORMATION:
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Goddard, Audrey
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Collin K.
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: P2930R1C4
 CURRENT APPLICATION NUMBER: US/10/033,396
 CURRENT FILING DATE: 2001-12-27
 PRIOR APPLICATION NUMBER: 60/095,325
 PRIOR FILING DATE: 1998-08-04
 PRIOR APPLICATION NUMBER: 60/112,851
 PRIOR FILING DATE: 1998-12-16
 PRIOR APPLICATION NUMBER: 60/113,145
 PRIOR FILING DATE: 1998-12-16
 PRIOR APPLICATION NUMBER: 60/113,511
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/115,558
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115,565
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115,733
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/119,341
 PRIOR FILING DATE: 1999-02-09
 PRIOR APPLICATION NUMBER: 60/119,537
 PRIOR FILING DATE: 1999-02-10
 PRIOR APPLICATION NUMBER: 60/119,965
 PRIOR FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: 60/162,506
 PRIOR FILING DATE: 1999-10-29
 PRIOR APPLICATION NUMBER: 60/170,262
 PRIOR FILING DATE: 1999-12-09
 PRIOR APPLICATION NUMBER: 60/187,202
 PRIOR FILING DATE: 2000-03-03
 PRIOR APPLICATION NUMBER: PCT/US99/12252
 PRIOR FILING DATE: 1999-06-02
 PRIOR APPLICATION NUMBER: PCT/US99/28634
 PRIOR FILING DATE: 1999-12-01
 PRIOR APPLICATION NUMBER: PCT/US99/28551
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US00/03565
 PRIOR FILING DATE: 2000-02-11
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: PCT/US00/05841
 PRIOR FILING DATE: 2000-03-02
 PRIOR APPLICATION NUMBER: PCT/US00/08439
 PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: PCT/US00/14941
 PRIOR FILING DATE: 2000-05-30
 PRIOR APPLICATION NUMBER: PCT/US00/15264
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: 2000-12-01
 NUMBER OF SEQ ID NOS: 38

SEQ ID NO 22
 LENGTH: 1029
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-033-396-22

Query Match 100.0%; Score 5241; DB 9; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MHILVHNAVYLLTLPGRADDSFQALDIMPPEKPLTAFLVDSSEALLPDLKTL 60
 1 MHILVHNAVYLLTLPGRADDSFQALDIMPPEKPLTAFLVDSSEALLPDLKTL 60
 61 RMIRSEVRLVDAADLEPQOLLFFVQSGIPVSSSKLLQFLDAQVAHDPQLEQNM 120
 61 RMIRSEVRLVDAADLEPQOLLFFVQSGIPVSSSKLLQFLDAQVAHDPQLEQNM 120
 121 DKNYMAHLVEYQHERGASGGQTFHSLTASLPFRDSTEARPKRSSPEQPIGGRIRVGT 180
 121 DKNYMAHLVEYQHERGASGGQTFHSLTASLPFRDSTEARPKRSSPEQPIGGRIRVGT 180
 181 QLRVLGPEDDLQAGMFLQIFPLSPDPMQSSPPRVALALQALQGLARVQGSPEVPGI 240
 181 QLRVLGPEDDLQAGMFLQIFPLSPDPMQSSPPRVALALQALQGLARVQGSPEVPGI 240
 241 TVRVQLALATLLSSPHGALVMSHRSHFLACPLLRQLCOYQRCVPDQGFSSLEKVL 300
 241 TVRVQLALATLLSSPHGALVMSHRSHFLACPLLRQLCOYQRCVPDQGFSSLEKVL 300
 301 QMLQWLDSPVEGGPRLAQLRMLASQASAGRISDVAGGILRLAFLAFLRQDLLEVASTV 360
 301 QMLQWLDSPVEGGPRLAQLRMLASQASAGRISDVAGGILRLAFLAFLRQDLLEVASTV 360
 361 RAVATLRSGECQVEEDLISKVLQGLIEVSPHLEELTAFFSATADAASPPFACKPVV 420
 361 RAVATLRSGECQVEEDLISKVLQGLIEVSPHLEELTAFFSATADAASPPFACKPVV 420
 421 VVSSLLEQEEEPPLAGKPGADGSLAVRGPSSGLLVDMLEMDPEVSSCPDLQRL 480
 421 VVSSLLEQEEEPPLAGKPGADGSLAVRGPSSGLLVDMLEMDPEVSSCPDLQRL 480
 481 FSRKRGGAQAVPSFRPYLLTLFTHOSWPTLHOCIFVLLGKSRQRPDSASIDFLMAC 540
 481 FSRKRGGAQAVPSFRPYLLTLFTHOSWPTLHOCIFVLLGKSRQRPDSASIDFLMAC 540
 541 IHVRIWQGRDQRTPOKRREELVLRVGPPELISVLELTAETRSODGDTAACSLIQAR 600
 541 IHVRIWQGRDQRTPOKRREELVLRVGPPELISVLELTAETRSODGDTAACSLIQAR 600
 601 LPLLSCCGDDESVRKVTEHLSCIQMGDSVIGRCRDLLOLYQRPRLRPVEVL 660
 601 LPLLSCCGDDESVRKVTEHLSCIQMGDSVIGRCRDLLOLYQRPRLRPVEVL 660
 661 LHSEGAASSVCYKLDGLIHRFTLLADTSDSRALENGADASMCKRLAVAHPLLRHL 720
 661 LHSEGAASSVCYKLDGLIHRFTLLADTSDSRALENGADASMCKRLAVAHPLLRHL 720
 721 PMIAALLHGRTHLNFQFROQNLSCFLHYVGLLELLQPHVFRSEHOGALMDCLSFIRL 780
 721 PMIAALLHGRTHLNFQFROQNLSCFLHYVGLLELLQPHVFRSEHOGALMDCLSFIRL 780
 781 LLNRYKSSRHLLAFINKFVGFIRKITYYNAFAISFLQKHADPHDLSFDSNDLVMLKSL 840
 781 LLNRYKSSRHLLAFINKFVGFIRKITYYNAFAISFLQKHADPHDLSFDSNDLVMLKSL 840
 841 LAGLSLSDRDRTRGDDEBEESASGSLPVVSILFTPLTAEMAPYMKRLSRGQTV 900
 841 LAGLSLSDRDRTRGDDEBEESASGSLPVVSILFTPLTAEMAPYMKRLSRGQTV 900
 901 DLEVLSDIDEMSRREILISFSTNLRQMLSSAECCRNLAFLALRSKMONSISIAAF 960
 901 DLEVLSDIDEMSRREILISFSTNLRQMLSSAECCRNLAFLALRSKMONSISIAAF 960

Qy 961 LPTFWKLSGSDPEFVQVOTALRNLPEYALLCOEHAVALHRAFLVGYGMDSQAQISEAL 1020
 |||||||
 Db 961 LPTFWKLSGSDPEFVQVOTALRNLPEYALLCOEHAVALHRAFLVGYGMDSQAQISEAL 1020
 |||||||
 Qy 1021 RILHMEAVM 1029
 |||||||
 Db 1021 RILHMEAVM 1029
 |||||||
 RESULT 10
 US-10-033-246-22
 ; Sequence 22, Application US/10033246
 ; Patent No. US20020098505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2930R1C12
 ; CURRENT FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: US/10/033,246
 ; PRIOR FILING DATE: 1998-08-04
 ; PRIOR APPLICATION NUMBER: 60/095,325
 ; PRIOR FILING DATE: 1998-08-04
 ; PRIOR APPLICATION NUMBER: 60/112,851
 ; PRIOR FILING DATE: 1998-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,145
 ; PRIOR FILING DATE: 1998-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,511
 ; PRIOR FILING DATE: 1998-12-22
 ; PRIOR APPLICATION NUMBER: 60/115,558
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/115,565
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/115,733
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/119,341
 ; PRIOR FILING DATE: 1999-02-09
 ; PRIOR APPLICATION NUMBER: 60/119,537
 ; PRIOR FILING DATE: 1999-02-10
 ; PRIOR APPLICATION NUMBER: 60/119,965
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: 60/162,506
 ; PRIOR FILING DATE: 1999-10-29
 ; PRIOR APPLICATION NUMBER: 60/170,262
 ; PRIOR FILING DATE: 1999-12-09
 ; PRIOR APPLICATION NUMBER: 60/187,202
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: PCT/US99/12252
 ; PRIOR FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28634
 ; PRIOR FILING DATE: 1999-12-01
 ; PRIOR APPLICATION NUMBER: PCT/US99/28551
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/03565
 ; PRIOR FILING DATE: 2000-02-11
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: PCT/US00/05841
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/08439
 ; PRIOR FILING DATE: 2000-03-30

;; PRIOR APPLICATION NUMBER: PCT/US00/14941
 ;; PRIOR FILING DATE: 2000-05-30
 ;; PRIOR APPLICATION NUMBER: PCT/US00/15264
 ;; PRIOR FILING DATE: 2000-06-02
 ;; PRIOR APPLICATION NUMBER: PCT/US00/32678
 ;; PRIOR FILING DATE: 2000-12-01
 ;; NUMBER OF SEQ ID NOS: 38
 ;; SEQ ID NO 22
 ;; LENGTH: 1029
 ;; TYPE: PRT
 ;; ORGANISM: Homo sapiens
 US-10-033-246-22
 Query Match 100.0%; Score 5241; DB 12; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MHLVHNAVILITLGPRAADSEFOALLDIWPEEKPPLPTAFVDTSEBALLPDMKL 60
 |||||||
 Db 1 MHLVHNAVILITLGPRAADSEFOALLDIWPEEKPPLPTAFVDTSEBALLPDMKL 60
 |||||||
 Qy 61 RMIRSEVLRVDAALODLEPQQLLFVOSGIPVSSMSKILQFLDQAVAHDPQTEONIM 120
 |||||||
 Db 61 RMIRSEVLRVDAALODLEPQQLLFVOSGIPVSSMSKILQFLDQAVAHDPQTEONIM 120
 |||||||
 Qy 121 DKNYMAHVEVQHERGASGGQTFHSLTASLPRRDSTEAPKPKSSPEQPIGGRIRVGT 180
 |||||||
 Db 121 DKNYMAHVEVQHERGASGGQTFHSLTASLPRRDSTEAPKPKSSPEQPIGGRIRVGT 180
 |||||||
 Qy 181 QLRVLGPEDDLQAGMFLQIFPLSPDPWQSSPPRVALAQALQGLARVQGSPEVPGI 240
 |||||||
 Db 181 QLRVLGPEDDLQAGMFLQIFPLSPDPWQSSPPRVALAQALQGLARVQGSPEVPGI 240
 |||||||
 Qy 241 TVRVQALATLTLSSPHGALVMSHRSHPFLACPLRQLOQYQRCVQDPTGFSSFLKVL 300
 |||||||
 Db 241 TVRVQALATLTLSSPHGALVMSHRSHPFLACPLRQLOQYQRCVQDPTGFSSFLKVL 300
 |||||||
 Qy 301 QMLQWLDSFVGEGLPRAQLRMLASQASAGRRLSDYRGGLIRLAELARFODLEVSVTV 360
 |||||||
 Db 301 QMLQWLDSFVGEGLPRAQLRMLASQASAGRRLSDYRGGLIRLAELARFODLEVSVTV 360
 |||||||
 Qy 361 RAVIATLRSGECCVPEPDLISKVLQGLIEYRSHBELTAFTSATADAAPPACKPVY 420
 |||||||
 Db 361 RAVIATLRSGECCVPEPDLISKVLQGLIEYRSHBELTAFTSATADAAPPACKPVY 420
 |||||||
 Qy 421 VVSSLLQEEPLAGGKPGADGSLAVRIGPSSGLVWLEMLDEPEVSSCPDLQRL 480
 |||||||
 Db 421 VVSSLLQEEPLAGGKPGADGSLAVRIGPSSGLVWLEMLDEPEVSSCPDLQRL 480
 |||||||
 Qy 481 FSRKRGQAQVPSFRPYLLTFTHOSWPTLHQCIRVLGKSREORFDPASLDFLMAC 540
 |||||||
 Db 481 FSRKRGQAQVPSFRPYLLTFTHOSWPTLHQCIRVLGKSREORFDPASLDFLMAC 540
 |||||||
 Qy 541 IHVPRIMQGRDQTPQKRREELVLRVQGPGLISLVELLAEATRSQDGGTAACSLQAR 600
 |||||||
 Db 541 IHVPRIMQGRDQTPQKRREELVLRVQGPGLISLVELLAEATRSQDGGTAACSLQAR 600
 |||||||
 Qy 601 IPLLSCCGGDDSVKRYTEHLSGCIQOMGDSVLGRCRDLQLQVLRQRELVPPEVL 660
 |||||||
 Db 601 IPLLSCCGGDDSVKRYTEHLSGCIQOMGDSVLGRCRDLQLQVLRQRELVPPEVL 660
 |||||||
 Qy 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASNAACKLVAHPDLLLRH 720
 |||||||
 Db 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASNAACKLVAHPDLLLRH 720
 |||||||
 Qy 721 PMTALLHGRTILNFOEFQONHLSCLFVLYGLLELLQVHVRSEHGALMDCLSFIRL 780
 |||||||
 Db 721 PMTALLHGRTILNFOEFQONHLSCLFVLYGLLELLQVHVRSEHGALMDCLSFIRL 780
 |||||||
 Qy 781 LNTYKSSRHIAFTIKFQOFHKYTYNAPAAISFLQKHADPLHLSFSDNSDLVWLKSL 840
 |||||||
 Db 781 LNTYKSSRHIAFTIKFQOFHKYTYNAPAAISFLQKHADPLHLSFSDNSDLVWLKSL 840
 |||||||

QY 841 LAGLSPPSRDRTDRGLDEGESESSAGSLPLVSVSLFTPLTAENAPYKKRLSGQTV 900
DB 841 LAGLSPPSRDRTDRGLDEGESESSAGSLPLVSVSLFTPLTAENAPYKKRLSGQTV 900
QY 901 DLLEVSIDIDEMRRRPELISFSTNQRMSAEECCRLALSLRSMQNSPSTAAAF 960
DB 901 DLLEVSIDIDEMRRRPELISFSTNQRMSAEECCRLALSLRSMQNSPSTAAAF 960
QY 961 LPTFMVCLGSQDEEVVOTALRNLPEYALLCOEHAVALHRAFLVGMVGMQDPSAQISEAL 1020
DB 961 LPTFMVCLGSQDEEVVOTALRNLPEYALLCOEHAVALHRAFLVGMVGMQDPSAQISEAL 1020
QY 1021 RLHMEAVM 1029
DB 1021 RLHMEAVM 1029
RESULT 11
US-10-033-301-22
Sequence 22, Application US/10033301
Patent No. US20020098506A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferreira, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same
FILE REFERENCE: F2930R1C6
CURRENT APPLICATION NUMBER: US/10/033,301
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565

QY 1 PRIOR FILING DATE: 2000-02-11
DB 1 PRIOR APPLICATION NUMBER: PCT/US00/04414
QY 1 PRIOR FILING DATE: 2000-02-22
DB 1 PRIOR APPLICATION NUMBER: PCT/US00/05841
QY 1 PRIOR FILING DATE: 2000-03-02
DB 1 PRIOR APPLICATION NUMBER: PCT/US00/08439
QY 1 PRIOR FILING DATE: 2000-03-30
DB 1 PRIOR APPLICATION NUMBER: PCT/US00/14941
QY 1 PRIOR FILING DATE: 2000-05-30
DB 1 PRIOR APPLICATION NUMBER: PCT/US00/15264
QY 1 PRIOR FILING DATE: 2000-06-02
DB 1 PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 1 PRIOR FILING DATE: 2000-12-01
DB 1 NUMBER OF SEQ ID NOS: 38
QY 1 SEQ ID NO: 22
DB 1 LENGTH: 1029
QY 1 TYPE: PRT
DB 1 ORGANISM: Homo sapiens
US-10-033-301-22
Query Match 100.0%; Score 5241; DB 12; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHILVHAMVITLTLGPPRADDSFQALDIWPEEKPLPTAFLVDTSEALLPMLKL 60
DB 1 MHILVHAMVITLTLGPPRADDSFQALDIWPEEKPLPTAFLVDTSEALLPMLKL 60
QY 61 RMIRSEVRLVDALDLEPOQLLFFVOSGIFVSSSKLQFLDQAVHPDQLEONIM 120
DB 61 RMIRSEVRLVDALDLEPOQLLFFVOSGIFVSSSKLQFLDQAVHPDQLEONIM 120
QY 121 DKNYMAHVEVQHERGASGGQTFHSLTASLPPRODTEAPKPKSSPEQPIGGRIRVGT 180
DB 121 DKNYMAHVEVQHERGASGGQTFHSLTASLPPRODTEAPKPKSSPEQPIGGRIRVGT 180
QY 181 QLRVLPEDDLACMFLOIFPLSPDPKRWSSPPRVALALQALQGLARVQSSPEVPGI 240
DB 181 QLRVLPEDDLACMFLOIFPLSPDPKRWSSPPRVALALQALQGLARVQSSPEVPGI 240
QY 241 TVRYLQALATLTLSSPHGALVMSMHRSHFLACPLLRQLCOYKCPQDTGFSSILFLKVL 300
DB 241 TVRYLQALATLTLSSPHGALVMSMHRSHFLACPLLRQLCOYKCPQDTGFSSILFLKVL 300
QY 301 QMLQWLDSPEVEGGPLRAQLMLASQASAGRRLSDVGRGLRLAEALAFQDLEVSSIV 360
DB 301 QMLQWLDSPEVEGGPLRAQLMLASQASAGRRLSDVGRGLRLAEALAFQDLEVSSIV 360
QY 361 RAVATLRSGEQSVEDLSKYVQGLIEVRSPLBELTAFSATDAASPACRPV 420
DB 361 RAVATLRSGEQSVEDLSKYVQGLIEVRSPLBELTAFSATDAASPACRPV 420
QY 421 VVSSLILQEEPEPLAGKPKGADGSLNAVRLGPSSGLVDMLEMDPEVSSCPDLQRL 480
DB 421 VVSSLILQEEPEPLAGKPKGADGSLNAVRLGPSSGLVDMLEMDPEVSSCPDLQRL 480
QY 481 FSRKKGQAQVPSFRYLLTLTHOSSWPTLMOCTIVLLGKSEQKQFDSASIDFLMAC 540
DB 481 FSRKKGQAQVPSFRYLLTLTHOSSWPTLMOCTIVLLGKSEQKQFDSASIDFLMAC 540
QY 541 IHVPRIMQGRDQRTPOKRRBELVLRVQGPETISVELILAEATRSQDGTACSLQAR 600
DB 541 IHVPRIMQGRDQRTPOKRRBELVLRVQGPETISVELILAEATRSQDGTACSLQAR 600
QY 601 LPLLLSCCGDDSVKRYTEHLSCGICQWQSDVLRRCRDLQLQYLQRELRVPVEVL 660
DB 601 LPLLLSCCGDDSVKRYTEHLSCGICQWQSDVLRRCRDLQLQYLQRELRVPVEVL 660
QY 661 LHSEGAASSVCKLDGLIHFTITLADTSSRALENAGAASMAKRLAVALHPLLLRL 720
DB 661 LHSEGAASSVCKLDGLIHFTITLADTSSRALENAGAASMAKRLAVALHPLLLRL 720
QY 721 PMIALLLHGRTHLNFQGRQNNHLSCLFVHVLGLLELLQPHVFRSEHGALMDCILSIFRL 780

|||||
Db 721 PMAIALHRTHLNFEQFEQONHLSCEFLVHLGELLELPHFVRSHEQALMDCILSFIRL 780
Qy 781 LLYNRSSSHLAATNKFFQFTHKYYTNAAPASIFLOKHADPLHLDNSDLVWLKSL 840
Db 781 LLYNRSSSHLAATNKFFQFTHKYYTNAAPASIFLOKHADPLHLDNSDLVWLKSL 840
Qy 841 LAGLSLPSHSDRTDRGLDEGEESAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQYVE 900
Db 841 LAGLSLPSHSDRTDRGLDEGEESAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQYVE 900
Qy 901 DLLEVLSDIDEMSRRRPELISFEFTNLQRLMSAECCCRNLAFSLARSMQNSPSIAAF 960
Db 901 DLLEVLSDIDEMSRRRPELISFEFTNLQRLMSAECCCRNLAFSLARSMQNSPSIAAF 960
Qy 961 LFFPMICLSQDEFEVQVOTLRNLPEYALLCOEHAVALLRAPLVGMVGMDSQAQISEAL 1020
Db 961 LFFPMICLSQDEFEVQVOTLRNLPEYALLCOEHAVALLRAPLVGMVGMDSQAQISEAL 1020
Qy 1021 RILHMEAVM 1029
Db 1021 RILHMEAVM 1029

RESULT 12
US-10-033-326-22
Sequence 22, Application US/10033326
Patent No. US20020098507A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Fond, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1C8
CURRENT APPLICATION NUMBER: US/10/033,326
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03

7/25/02

PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 22
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-10-033-326-22
Query Match 100.0%; Score 5241; DB 12; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHLLVYHAWYILTLTGPRRADSEFOALDIMEPEKPLPTAFVPTSEALLPWLK 60
Db 1 MHLLVYHAWYILTLTGPRRADSEFOALDIMEPEKPLPTAFVPTSEALLPWLK 60
Qy 61 RMIRSEVRLVDAALDLEPQOLLFLVQSGFIPVSSMSKLLQFLDAVAHDPTLEQNM 120
Db 61 RMIRSEVRLVDAALDLEPQOLLFLVQSGFIPVSSMSKLLQFLDAVAHDPTLEQNM 120
Qy 121 DKNYMAHVEVCHERASGGQTFHSLTASLPRRRSTEARPKRSSPEQITGGRIYGT 180
Db 121 DKNYMAHVEVCHERASGGQTFHSLTASLPRRRSTEARPKRSSPEQITGGRIYGT 180
Qy 181 QLRVLGPEDDLAQMFQIFLSPDPWQSSPPRVALALQALGOELARVQSPVEYGI 240
Db 181 QLRVLGPEDDLAQMFQIFLSPDPWQSSPPRVALALQALGOELARVQSPVEYGI 240
Qy 241 TVRVQALATLSSPHGALVMSHNSHTACPLRLQCOYQRCVPDITGFSLSFLKYL 300
Db 241 TVRVQALATLSSPHGALVMSHNSHTACPLRLQCOYQRCVPDITGFSLSFLKYL 300
Qy 301 QMLQMLDPSVEGGLRAOLRLASQASAGRRLSDVRGILLRLAEALAFQDLEVVSTV 360
Db 301 QMLQMLDPSVEGGLRAOLRLASQASAGRRLSDVRGILLRLAEALAFQDLEVVSTV 360
Qy 361 RAVIATLRSGEQSVPEPDLISKVLQGLEVRSPHLEELTAFSAADAAAPPACKPVY 420
Db 361 RAVIATLRSGEQSVPEPDLISKVLQGLEVRSPHLEELTAFSAADAAAPPACKPVY 420
Qy 421 VVSSLLQEEEPPLAGKPGADGSLAEAVRLGPSSGLVDMLEMLDEPVSSCDLQRL 480
Db 421 VVSSLLQEEEPPLAGKPGADGSLAEAVRLGPSSGLVDMLEMLDEPVSSCDLQRL 480
Qy 481 FSRKRGGAQVPSFRPYLLTLTHOSSWPTLHOCIRVILIGKREORFQDSALDFLMAC 540
Db 481 FSRKRGGAQVPSFRPYLLTLTHOSSWPTLHOCIRVILIGKREORFQDSALDFLMAC 540
Qy 541 IHVPRIWGRDQTPQKRREELVRYOGPELISVELLILAEATRSQDGTAACSLIQAR 600
Db 541 IHVPRIWGRDQTPQKRREELVRYOGPELISVELLILAEATRSQDGTAACSLIQAR 600
Qy 601 LPLLSCCGDDESVRKVTBHLSGCIQMGDSVLGRRCDLLQLYLQRELPVPEVL 660
Db 601 LPLLSCCGDDESVRKVTBHLSGCIQMGDSVLGRRCDLLQLYLQRELPVPEVL 660

Db 601 LPLLSCCGDDESVKRYTEHLSCGICQMGDSVIGRRCDLLDLYLRPELRVPVPEVL 660
Qy 661 LHSEGAASSVCKLDGLIHRFTITLADTSDSRALENRGADASMACRKLAVAHPLLLRL 720
Db 661 LHSEGAASSVCKLDGLIHRFTITLADTSDSRALENRGADASMACRKLAVAHPLLLRL 720
Qy 721 PMIAALLHGRTHLNFQEFROQNLSCFLHVLGLLELLQPHVFRSEHOGALMDCLLSFRL 780
Db 721 PMIAALLHGRTHLNFQEFROQNLSCFLHVLGLLELLQPHVFRSEHOGALMDCLLSFRL 780
Qy 781 LLNKRSRHLAAFINKFVQFIHKYITYNAPAAISFLQKHADPLHDLSFQNSDLVMLKSL 840
Db 781 LLNKRSRHLAAFINKFVQFIHKYITYNAPAAISFLQKHADPLHDLSFQNSDLVMLKSL 840
Qy 841 LAGLSRSDRTDRGLDEEGEESAGSLPLVSVSLFTPLTAEMAPYMKRLSNGQVTE 900
Db 841 LAGLSRSDRTDRGLDEEGEESAGSLPLVSVSLFTPLTAEMAPYMKRLSNGQVTE 900
Qy 901 DLEVLSDIDEMSRRPETLSFSTNQLRLMSAECCRNIAFSLARSMONSPSIAAF 960
Db 901 DLEVLSDIDEMSRRPETLSFSTNQLRLMSAECCRNIAFSLARSMONSPSIAAF 960
Qy 961 LPTMYCLGSODEFVQVOTALRNLPYALCOEHAAVLHRAFLVGMGOMPQSAQISAL 1020
Db 961 LPTMYCLGSODEFVQVOTALRNLPYALCOEHAAVLHRAFLVGMGOMPQSAQISAL 1020
Qy 1021 RLHMEAVM 1029
Db 1021 RLHMEAVM 1029

RESULT 13
US-09-984-271-235
; Sequence 235, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482, 273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092, 921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092, 922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092, 956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 235
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-271-235

Query Match 39.2%; Score 205; DB 9; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.7e-157;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 626 IQQMGDSVIGRRCDLLDLYLRPELRVPVPEVLHSEGAASSVCKLDGLIHRFTITL 685
Db 1 IQQMGDSVIGRRCDLLDLYLRPELRVPVPEVLHSEGAASSVCKLDGLIHRFTITL 60
Qy 686 ADTSDSRALENRGADASMACRKLAVAHPLLLRLPMIAALLHGRTHLNFQEFROQNL 745
Db 61 ADTSDSRALENRGADASMACRKLAVAHPLLLRLPMIAALLHGRTHLNFQEFROQNL 120
Qy 746 CFLHVLGLLELLQPHVFRSEHOGALMDCLLSFIRLLNKRSSRHLAAFINKFVQFIHKY 805
Db 746 CFLHVLGLLELLQPHVFRSEHOGALMDCLLSFIRLLNKRSSRHLAAFINKFVQFIHKY 805

Db 121 CFLHVLGLLELLQPHVFRSEHOGALMDCLLSFIRLLNKRSSRHLAAFINKFVQFIHKY 180
Qy 806 ITYNAPAAISFLQKHADPLHDLSFQNSDLVMLKSLAGLSRSDRTDRGLDEEGEES 865
Db 181 ITYNAPAAISFLQKHADPLHDLSFQNSDLVMLKSLAGLSRSDRTDRGLDEEGEES 240
Qy 866 SAGSLPLVSVSLFTPLTAEMAPYMKRLSNGQVTELLVSDIDEMSRRPETLSFST 925
Db 241 SAGSLPLVSVSLFTPLTAEMAPYMKRLSNGQVTELLVSDIDEMSRRPETLSFST 300
Qy 926 NLQRLMSAECCRNIAFSLARSMONSPSIAAFPLPFMYCLGSODEFVQVOTALRNLP 985
Db 301 NLQRLMSAECCRNIAFSLARSMONSPSIAAFPLPFMYCLGSODEFVQVOTALRNLP 360
Qy 986 YALCOEHAAVLHRAFLVGMGOMPQSAQISALRLHMEAVM 1029
Db 361 YALCOEHAAVLHRAFLVGMGOMPQSAQISALRLHMEAVM 404

RESULT 14
US-09-984-271-148
; Sequence 148, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482, 273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092, 921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092, 922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092, 956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 148
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (328)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-148

Query Match 31.8%; Score 165; DB 9; Length 328;
Best Local Similarity 100.0%; Pred. No. 9.5e-126;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 703 MACRKLAVAHPLLLRLPMIAALLHGRTHLNFQEFROQNLSCFLHVLGLLELLQPHV 762
Db 1 MACRKLAVAHPLLLRLPMIAALLHGRTHLNFQEFROQNLSCFLHVLGLLELLQPHV 60
Qy 763 RSEHOGALMDCLLSFIRLLNKRSSRHLAAFINKFVQFIHKYITYNAPAAISFLQKHAD 822
Db 61 RSEHOGALMDCLLSFIRLLNKRSSRHLAAFINKFVQFIHKYITYNAPAAISFLQKHAD 120
Qy 823 PLHDLSPFQNSDLVMLKSLAGLSRSDRTDRGLDEEGEESAGSLPLVSVSLFTPLT 882
Db 121 PLHDLSPFQNSDLVMLKSLAGLSRSDRTDRGLDEEGEESAGSLPLVSVSLFTPLT 180
Qy 883 AAEMAPYMKRLSNGQVTELLVSDIDEMSRRPETLSFSTNLQRLMSAECCRNLA 942
Db 181 AAEMAPYMKRLSNGQVTELLVSDIDEMSRRPETLSFSTNLQRLMSAECCRNLA 240
Qy 943 FSLARSMONSPSIAAFPLPFMYCLGSODEFVQVOTALRNLPYALCOEHAAVLHRAFL 1002
Db 943 FSLARSMONSPSIAAFPLPFMYCLGSODEFVQVOTALRNLPYALCOEHAAVLHRAFL 1002

Db 241 FSLALRSMQNSPSIAAFLPTFMYCIGSODFEVQTALENLPEYALCOEHAAVLHRAAF 300
QY 1003 LVGMTGMDPSAQSISEALRIHMEAVM 1029
Db 301 LVGMTGMDPSAQSISEALRIHMEAVM 327

RESULT 15
US-09-764-891-4155
; Sequence 4155, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4155
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (40)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids.
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (94)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (98)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (139)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (141)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4155

Query Match 10.1%; Score 530.5; DB 9; Length 161;
Best Local Similarity 73.1%; Pred. No. 5.5e-35;
Matches 117; Conservative 7; Mismatches 17; Indels 19; Gaps 3;

QY 812 AAISFQKADPLADLSPDSDLVMLKSLAGLSLPSRDDRTDGLDEGEESAGSLP 871
Db 2 AAISFQKADPLADLSPDSDLVMLKSLAGLSLPSRKDRTDGLDEGEESAGSLP 61
QY 872 LVSVSLFTPLTAEMAPYMKRLSRGQTV-----DLLEVLSDIDEMSRRPEIIS 921
|||||

Db 62 LVSVSLFTPLTAEMAPYMKRLXRGQTVGESXPXSTPDLLLEXLSDIDEMSRRPEIIS 121
QY 922 FFSINLQRLMSAEE-----CCRNLAFLSLRSMQNSPSI 956
Db 122 FFFDQ-----PAADELURPXKCPXLPALNSCRTPAL 157
|||

Search completed: May 7, 2003, 19:43:42
Job time : 33 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:37:58 ; Search time 44 Seconds

(Without alignments)
3116.247 Million cell updates/sec

Title: US-09-929-769-7

Perfect score: 5241
Sequence: 1 MHILVYHAMVILLTLCPPRA.....MDPSAQISEALRIHMEAVM 1029

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

1: A.Geneseq_101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5241	100.0	1029	21	Human PRO3434 prot
2	5241	100.0	1029	21	PRO3434, a novel s
3	5241	100.0	1029	23	Tumour associated
4	3229	61.6	650	21	Human ORFX ORF2259
5	2055	39.2	404	21	Human gene 46 enco
6	2055	39.2	404	22	Human gene 46 enco
7	1665	31.8	404	23	Human gene 46 enco
8	1665	31.8	327	22	Human gene 46 enco
9	1665	31.8	327	23	Human gene 46 enco
10	1665	31.8	328	21	Human secreted pro

11	863	16.5	2042	22	AB859689
12	530.5	10.1	161	22	AAW5497
13	156	3.0	2472	22	AAU33568
14	149	2.8	1217	22	AAW52358
15	143	2.7	1519	21	AAW85650
16	140	2.7	1778	22	AAW79480
17	140	2.7	2099	22	AAW78496
18	138	2.6	3674	22	AAU14347
19	137.5	2.6	2757	22	ABG09921
20	135	2.6	2014	22	AAW67395
21	135	2.6	2014	23	AAE24137
22	133.5	2.5	931	22	AAW39815
23	133	2.5	737	15	AAW5494
24	133	2.5	737	17	AAW60804
25	133	2.5	737	18	AAW5019
26	132	2.5	2444	22	ABW62786
27	131	2.5	1302	22	AAW00761
28	128	2.4	1339	22	ABW21500
29	128	2.4	2473	22	AAW78911
30	127.5	2.4	3353	23	AAU82706
31	127	2.4	2733	22	ABW60737
32	127	2.4	3338	22	ABG09919
33	126	2.4	745	22	AAW41083
34	126	2.4	878	19	AAW54373
35	126	2.4	878	20	AAW5397
36	126	2.4	878	22	AAW5393
37	126	2.4	879	22	AAW31635
38	126	2.4	881	22	AAW31634
39	125.5	2.4	2779	22	ABW62371
40	125	2.4	901	22	AAU1540
41	125	2.4	2116	22	ABW58110
42	125	2.4	2383	23	ABW56531
43	125	2.4	2668	22	AAW79895
44	125	2.4	2797	22	AAU31359
45	125	2.4	3007	22	AAU28053

ALIGNMENTS

RESULT 1	
AAW24027	standard; Protein; 1029 AA.
ID	AAW24027
XX	AAW24027
XX	25-JAN-2001 (first entry)
DT	Human PRO3434 protein sequence SEQ ID NO:23.
XX	Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
DE	Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
XX	Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
KW	Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
KW	Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
XX	Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
OS	Homo sapiens.
XX	MO200053750-A1.
PN	14-SEP-2000.
PD	14-SEP-2000.
XX	02-DEC-1999; 99WO-US28551.
PF	02-DEC-1999; 99WO-US28551.
XX	08-MAR-1999; 99WO-US05028.
PR	01-SEP-1999; 99WO-US05011.
PR	29-OCT-1999; 99US-0162506.
PR	30-NOV-1999; 99WO-US28313.
PR	01-DEC-1999; 99WO-US28634.
XX	(GERTH) GENENTECH INC.
PA	Bozstein D, Goddard A, Gurney AL, Roy MA, Watanabe CK, Wood WL;
PI	Bozstein D, Goddard A, Gurney AL, Roy MA, Watanabe CK, Wood WL;
XX	WPL: 2000-594320/56.
DR	N-PSDB: AAC58109.

XX Antibodies specific for PRO polypeptides, used to diagnose and inhibit
 PT the growth of tumors in mammals, and to identify inhibitors of PRO
 PT polypeptide activity or expression -
 PS
 PS Claim 61; Fig 14; 226pp; English.
 CC The present invention describes an antibody that binds to a human
 CC protein (I) selected from: PRO381; PRO1269; PRO1410; PRO155; PRO1760;
 CC PRO3434; PRO1293; PRO1393; PRO1303; PRO344; PRO4354;
 CC PRO4397; PRO1555; PRO1096; PRO2038; and PRO2262. (I) has
 CC anticancer activity and can be used to diagnose tumors in mammals, by
 CC detecting complex formation when the antibody is contacted with test
 CC cells. Increased expression of genes encoding (I) can also be detected
 CC to diagnose tumors. Agents which inhibit the activity of (I),
 CC especially the antibodies, or an antisense oligonucleotide which
 CC hybridises to genes encoding (I), can be used to inhibit tumor growth,
 CC preferably by inducing cell death. Methods from the present invention
 CC can be used to identify compounds which inhibit the biological activity
 CC of (I). AAC58019 to AAC58102 represent PCR primers and hybridisation
 CC probes used in examples from the present invention for human PRO
 CC sequences. AAC58103 to AAC58122 and AAB24021 to AAB24040 represent human
 CC PRO polynucleotide and protein sequences given in the exemplification of
 CC the present invention.
 CC
 XX Sequence 1029 AA:

Query Match 100.0%; Score 5241; DB:21; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHLLVHAAYVILLTGPAPADSEFOALDINPEEKPLPAVLTSEALLPWLK 60
 DB 1 MHLLVHAAYVILLTGPAPADSEFOALDINPEEKPLPAVLTSEALLPWLK 60
 QY 61 RMIRSEVLRLVDAALDLEPOOLLFVOSFGIVSSMSKLLQPLDOAVAHDPOTLEONIM 120
 DB 61 RMIRSEVLRLVDAALDLEPOOLLFVOSFGIVSSMSKLLQPLDOAVAHDPOTLEONIM 120
 QY 121 DKNYMAHLEVOHERGASGGQTFHSLTASLPPRDSTEAPKPKSSPEQIGGRIKRVCT 180
 DB 121 DKNYMAHLEVOHERGASGGQTFHSLTASLPPRDSTEAPKPKSSPEQIGGRIKRVCT 180
 QY 181 QLRVLGPEEDLAGMTQITPPLSPDPRWSSPPRVALALQALGOLARVQSSPEVPGI 240
 DB 181 QLRVLGPEEDLAGMTQITPPLSPDPRWSSPPRVALALQALGOLARVQSSPEVPGI 240
 QY 241 TVRVLOALATLSSPHGALVMSMRSHFLACPLRLQOLCOYQRCVQDITGESSLFLKVL 300
 DB 241 TVRVLOALATLSSPHGALVMSMRSHFLACPLRLQOLCOYQRCVQDITGESSLFLKVL 300
 QY 301 QMLQWLDSFGVEGGPLRAQLRMLASQASAGRRLSDVRGGLRLAELAFRODLEVVSVTV 360
 DB 301 QMLQWLDSFGVEGGPLRAQLRMLASQASAGRRLSDVRGGLRLAELAFRODLEVVSVTV 360
 QY 361 RAVIATLRGEGCSEVPDLISKVLOGLIEVRSHLELLTAFFSATAADASPAPACKPVV 420
 DB 361 RAVIATLRGEGCSEVPDLISKVLOGLIEVRSHLELLTAFFSATAADASPAPACKPVV 420
 QY 421 VVSSLLOEEEPPLAGKPGADGSLFAVRLGSSGLVLMLEMLDEVVSSCPDLQRL 480
 DB 421 VVSSLLOEEEPPLAGKPGADGSLFAVRLGSSGLVLMLEMLDEVVSSCPDLQRL 480
 QY 481 FSRRRKGAQVPSFRPYLLTFTHOSSWPTLHQCIRVLGKRGQRFDPASLDFLMAC 540
 DB 481 FSRRRKGAQVPSFRPYLLTFTHOSSWPTLHQCIRVLGKRGQRFDPASLDFLMAC 540
 QY 541 IIVPRITWGRDQRTPOKRRREELVLRVQGPBLSLVELLAAEATRSQDDDTAACSILQAR 600
 DB 541 IIVPRITWGRDQRTPOKRRREELVLRVQGPBLSLVELLAAEATRSQDDDTAACSILQAR 600
 QY 601 LPLLLSCCGDDESVRKVEHLSCGIQWGDVSLGRRCRDLQLQLPRLVAPPEVL 660
 DB 601 LPLLLSCCGDDESVRKVEHLSCGIQWGDVSLGRRCRDLQLQLPRLVAPPEVL 660

DB 601 LPLLLSCCGDDESVRKVEHLSCGIQWGDVSLGRRCRDLQLQLPRLVAPPEVL 660
 QY 661 LHSEGAASSVCKLDGLIHFRITLADTSDSRALENRGADASMACKRLAVAHPLLLRLH 720
 DB 661 LHSEGAASSVCKLDGLIHFRITLADTSDSRALENRGADASMACKRLAVAHPLLLRLH 720
 QY 721 PMTALLHGRTHLNFOEFROQNLSCFLHVLGLLELLQPHVFSEHQALMDCLLSFIRL 780
 DB 721 PMTALLHGRTHLNFOEFROQNLSCFLHVLGLLELLQPHVFSEHQALMDCLLSFIRL 780
 QY 781 LNLVRRSSRLAFIRKPFQFIKITYYNAAPASIFLOKHADPLHLSFQNSDLVLMKSL 840
 DB 781 LNLVRRSSRLAFIRKPFQFIKITYYNAAPASIFLOKHADPLHLSFQNSDLVLMKSL 840
 QY 841 LAGLSLPSRDRDTRGDEGESESSAGSLPVSVLFTPLTAENAPYVKRLSRQGYE 900
 DB 841 LAGLSLPSRDRDTRGDEGESESSAGSLPVSVLFTPLTAENAPYVKRLSRQGYE 900
 QY 901 DLLEVLSDIDEMSRRPETLISFSTNLQRLUMSSAECCRNLAFLSLRSMQNSPSTAAAF 960
 DB 901 DLLEVLSDIDEMSRRPETLISFSTNLQRLUMSSAECCRNLAFLSLRSMQNSPSTAAAF 960
 QY 961 LPTFMYCLGSDPEVYOTALRNLPEYALLCOEHAVALHRAFLVGYMGOMDPSAQISEAL 1020
 DB 961 LPTFMYCLGSDPEVYOTALRNLPEYALLCOEHAVALHRAFLVGYMGOMDPSAQISEAL 1020
 QY 1021 RILHMEAVM 1029
 DB 1021 RILHMEAVM 1029

RESULT 2
 AAY96736
 ID AAY96736 standard; Protein; 1029 AA.
 XX
 AC AAY96736;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE PRO3434, a novel secreted protein.
 XX
 KW PRO3434; secreted protein; transmembrane protein; recombinant production;
 KM gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT 1..16
 FT /label= Signal_peptide
 FT 91..97
 FT /note= "N-myristoylation site"
 FT 136..142
 FT /note= "N-myristoylation site"
 FT 154..158
 FT /note= "CAMP- and GMP-dependent protein kinase
 FT phosphorylation site"
 FT 224..230
 FT /note= "N-myristoylation site"
 FT 329..333
 FT /note= "amidation site"
 FT 331..335
 FT /note= "CAMP- and GMP-dependent protein kinase
 FT phosphorylation site"
 FT 435..441
 FT /note= "N-myristoylation site"
 FT 439..445
 FT /note= "N-myristoylation site"
 FT 443..449
 FT /note= "N-myristoylation site"
 FT 616..620
 FT /note= "CAMP- and GMP-dependent protein kinase
 FT phosphorylation site"
 FT Modified-site 634..636

FT /note="amidation site"
 FT Modified-site 665..671
 FT /note="N-myristoylation site"
 FT Modified-site 698..704
 FT /note="N-myristoylation site"
 FT Modified-site 785..789
 FT /note="cAMP- and cGMP-dependent protein kinase phosphorylation site"
 FT Modified-site 891..895
 FT /note="cAMP- and cGMP-dependent protein kinase phosphorylation site"
 PN MO200036102-A2.
 XX 22-JUN-2000.
 XX 01-DEC-1999; 99WO-US28634.
 XX 16-DEC-1998; 98US-0112851.
 PR 16-DEC-1998; 98US-0113145.
 PR 22-DEC-1998; 98US-0113511.
 PR 12-JAN-1999; 99US-0115558.
 PR 12-JAN-1999; 99US-0115565.
 PR 12-JAN-1999; 99US-0115733.
 PR 09-FEB-1999; 99US-0119341.
 PR 10-FEB-1999; 99US-0119537.
 PR 12-FEB-1999; 99US-0119965.
 PR 02-JUN-1999; 99WO-US12252.
 XX (GETH) GENENTECH INC.
 PI Botstein D, Desnoyers L, Ferrara N, Fong S, Gao W, Goddard A;
 PI Guney AL, Pan J, Roy MA, Stewart TA, Tumas D, Watanabe CK;
 PI Wood WL;
 XX WPI: 2000-431586/37.
 DR N-PSDB: AAA51266.
 XX
 PT Isolated nucleic acid molecule encodes a PRO polypeptide which is a
 PT transmembrane polypeptide
 PS Claim 12; Fig 16; 154pp: English.
 XX
 CC This is PRO3434, a novel secreted protein. The invention
 CC concerns novel secreted and transmembrane proteins, designated PRO.
 CC polypeptides. The cDNA and gene sequences are useful in the recombinant
 CC production of PRO polypeptides, as a hybridization probe to screen
 CC libraries to isolate cDNAs with sequence identity to PRO polypeptides or
 CC to map the gene encoding the PRO polypeptides and analyzing genetic
 CC disorders. The cDNA/gene can also be used to produce transgenic animals
 CC useful for the development and screening of therapeutically useful
 CC reagents. They can also be used in gene therapy, e.g. to replace a
 CC defective gene.
 CC
 XX
 SQ Sequence 1029 AA:
 Query Match 100.0%; Score 5241; DB 21; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 181 QLRVLGPPEDDLAQMFLQIFPLSPDPRMOSSSPRPVALALQALGDELARVVOGSEPEVPGI 240
 QY 241 TVRVQLATATLLSPHGALVWMSHRSHFLACPLLRLQCLQYORCPYOTGSSFLKYL 300
 DB 241 TVRVQLATATLLSPHGALVWMSHRSHFLACPLLRLQCLQYORCPYOTGSSFLKYL 300
 QY 301 QMLQWLDSPVGGGLRLQRLMLASQASAGRSLSPVGGGLRLALALAFRODLEVVSTV 360
 DB 301 QMLQWLDSPVGGGLRLQRLMLASQASAGRSLSPVGGGLRLALALAFRODLEVVSTV 360
 QY 361 RAVIATLRSGECCSVPEPDLISKVLQGLIEVRSPIHEELLTAFFSATDAASPFPACKPV 420
 DB 361 RAVIATLRSGECCSVPEPDLISKVLQGLIEVRSPIHEELLTAFFSATDAASPFPACKPV 420
 QY 421 VVSSLLQDEEPEPLAGKCGADGSGLEAVRLGSSGLVDYMLDPEVYSSCPDLQLRL 480
 DB 421 VVSSLLQDEEPEPLAGKCGADGSGLEAVRLGSSGLVDYMLDPEVYSSCPDLQLRL 480
 QY 481 FSRKRGKGOAQPSPRPYLLTFTHQSSWPTLHQCIRVLLGKSRBORFDPASLDPLMAC 540
 DB 481 FSRKRGKGOAQPSPRPYLLTFTHQSSWPTLHQCIRVLLGKSRBORFDPASLDPLMAC 540
 QY 541 IHVPRIMQGRDQRTPOKRRREELVLRVGGPELISVELLAEATRSQDGDPAACSLIOAR 600
 DB 541 IHVPRIMQGRDQRTPOKRRREELVLRVGGPELISVELLAEATRSQDGDPAACSLIOAR 600
 QY 601 LPLLSCCGDDESVRYKTEHLSCIQMGDSVLRRCRDLLQYLQRPRLRVVPEVL 660
 DB 601 LPLLSCCGDDESVRYKTEHLSCIQMGDSVLRRCRDLLQYLQRPRLRVVPEVL 660
 QY 661 LHSEGAASSVYCKLKGILHREFTLLADTSDSRALENRGADASMACRKLAVAHPLLLRHL 720
 DB 661 LHSEGAASSVYCKLKGILHREFTLLADTSDSRALENRGADASMACRKLAVAHPLLLRHL 720
 QY 721 PMIAALLHGRTHLNFQERROONHLSCLFHLVGLLELLOPHVFRSEHOGALMDCLLSFTRL 780
 DB 721 PMIAALLHGRTHLNFQERROONHLSCLFHLVGLLELLOPHVFRSEHOGALMDCLLSFTRL 780
 QY 781 LINTYKSSRHILAAFLINKFQFIHKYITYNAPALISFLQKHADPLDHSFNDNSDLVWLSSL 840
 DB 781 LINTYKSSRHILAAFLINKFQFIHKYITYNAPALISFLQKHADPLDHSFNDNSDLVWLSSL 840
 QY 841 LAGLSLPSRDKFTDGLDEEGEESSAGSLPLVSVSLFTPLTAAMAPYMKRLSGQTV 900
 DB 841 LAGLSLPSRDKFTDGLDEEGEESSAGSLPLVSVSLFTPLTAAMAPYMKRLSGQTV 900
 QY 901 DLLEVLSDIDEMSRRPETILSFSTNLQRLMSABECCRNLAFLSLALRSMONSPSIAAF 960
 DB 901 DLLEVLSDIDEMSRRPETILSFSTNLQRLMSABECCRNLAFLSLALRSMONSPSIAAF 960
 QY 961 LPTFMYCIGSQDFEVVQVATRLNLPETALICQHAHVLLHRAFLVGMGMDSQAISEAL 1020
 DB 961 LPTFMYCIGSQDFEVVQVATRLNLPETALICQHAHVLLHRAFLVGMGMDSQAISEAL 1020
 QY 1021 RLHMEAYM 1029
 DB 1021 RLHMEAYM 1029
 RESULT 3
 ID AU077765
 ID AU077765 standard; Protein: 1029 AA.
 XX AU077765;
 DE 05-JUN-2002 (first entry)
 XX Tumour associated antigenic target polypeptide (TAT) 156.
 XX Tumour associated antigenic target polypeptide: TAT; cancer:
 KW breast cancer; colorectal cancer; lung cancer; ovarian cancer;
 KW central nervous system cancer; liver cancer; bladder cancer;

Query Match	100.0%	Score 5241	DB 23	Length 1029
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1029	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1 MHILVAMVITLLTGPPRADSEFOALLDIMEPEKPLPTFAVLDTSEEAALLPDMTKL 60				
Db 1 MHILVAMVITLLTGPPRADSEFOALLDIMEPEKPLPTFAVLDTSEEAALLPDMTKL 60				
QY 61 RMISEVRLVDAALDLEPQULLFVQSGFIPVSSMSKLOFLDQAAVADPQTLEONIM 120				
Db 61 RMISEVRLVDAALDLEPQULLFVQSGFIPVSSMSKLOFLDQAAVADPQTLEONIM 120				
QY 121 DKNMHALVEQVHERGASGGQTFHSLTLTASLPBRDSTEARPKSSPDQIGQGRIRVGT 180				
Db 121 DKNMHALVEQVHERGASGGQTFHSLTLTASLPBRDSTEARPKSSPDQIGQGRIRVGT 180				
QY 181 QLRVLGDEDDLAGMFLQIFPLSPDPWQSSSRPVALAQQALGQELARVYQVSGPEVNGI 240				
Db 181 QLRVLGDEDDLAGMFLQIFPLSPDPWQSSSRPVALAQQALGQELARVYQVSGPEVNGI 240				
QY 241 TVRVLQALATLLSSPHRGALVMSMHSHTLACPLRLQCLQYRCYVPDTGSSFLKAYLL 300				
Db 241 TVRVLQALATLLSSPHRGALVMSMHSHTLACPLRLQCLQYRCYVPDTGSSFLKAYLL 300				
QY 301 QMLQMDLSPVEGGPLRAQLRLMASQASAGRISDVRGGLRLAEALAFRODLEVSVSTV 360				
Db 301 QMLQMDLSPVEGGPLRAQLRLMASQASAGRISDVRGGLRLAEALAFRODLEVSVSTV 360				
QY 361 RAVIATLRSEGGQSVEDPLISKVLQGLIEVRSPHLEELLTAFFSXTADAAEPFPAKPEVY 420				
Db 361 RAVIATLRSEGGQSVEDPLISKVLQGLIEVRSPHLEELLTAFFSXTADAAEPFPAKPEVY 420				
QY 421 VVSSLILQEEEPPLAGKPGADGSLFAVRLGSSGLVQWIMLEMDPEVYVSSCPDQLRL 480				
Db 421 VVSSLILQEEEPPLAGKPGADGSLFAVRLGSSGLVQWIMLEMDPEVYVSSCPDQLRL 480				
QY 481 FSRKRGGAQVPSFRPLYLTLLTFHSSWPTLHOCIRVLKRSQRQRPDSALDFLWAC 540				
Db 481 FSRKRGGAQVPSFRPLYLTLLTFHSSWPTLHOCIRVLKRSQRQRPDSALDFLWAC 540				
QY 541 ITHVRIMQGRDQTPQKRREELVLRVQSGELLISYVELLIAAEKETSQOGDTAAACSLIQAR 600				
Db 541 ITHVRIMQGRDQTPQKRREELVLRVQSGELLISYVELLIAAEKETSQOGDTAAACSLIQAR 600				
QY 601 LPLLSCCCDDDSVRYVTHLSGCIQQMGDSVLRGRCDLLDQLQYLRPELRAVPEVYL 660				
Db 601 LPLLSCCCDDDSVRYVTHLSGCIQQMGDSVLRGRCDLLDQLQYLRPELRAVPEVYL 660				
QY 661 LHSBGAASSVYCKLDGLIHRFITLLADTSDSRALENRGADASMACRKLAVAHPLLLRL 720				
Db 661 LHSBGAASSVYCKLDGLIHRFITLLADTSDSRALENRGADASMACRKLAVAHPLLLRL 720				
QY 721 PMTALLHGRTHLNPOEFHQNHLSLFLAVLGLELQPHVYRSHQCALMDCLLSTFRL 780				
Db 721 PMTALLHGRTHLNPOEFHQNHLSLFLAVLGLELQPHVYRSHQCALMDCLLSTFRL 780				

QY 781 L L N R K S S R H L A F I N K F V O F I H K Y I T Y N A P A I S F L O K H A D P L H D S F D N S D L V M L K S L 840
 DB 781 L L N R K S S R H L A F I N K F V O F I H K Y I T Y N A P A I S F L O K H A D P L H D S F D N S D L V M L K S L 840
 QY 841 L A G I S L S R D R T R G D E E G E E S S A G S L P L V S V S L F T P L T A A E M A P Y M K R L S R G Q I V E 900
 DB 841 L A G I S L S R D R T R G D E E G E E S S A G S L P L V S V S L F T P L T A A E M A P Y M K R L S R G Q I V E 900
 QY 901 D L E V L S D I D E M S R R P E I I S F F S T N I O R L M S S A E C C R N I A F S L A R S M O N S P S I A A F 960
 DB 901 D L E V L S D I D E M S R R P E I I S F F S T N I O R L M S S A E C C R N I A F S L A R S M O N S P S I A A F 960
 QY 961 L P T E M Y C L G S O D F E V Y O T A L R N L P E Y A L L C Q E H A V I L L H R A F L V G M Y G O M P S A Q I S E A L 1020
 DB 961 L P T E M Y C L G S O D F E V Y O T A L R N L P E Y A L L C Q E H A V I L L H R A F L V G M Y G O M P S A Q I S E A L 1020
 QY 1021 R I L H M E A V M 1029
 DB 1021 R I L H M E A V M 1029

RESULT 4

AAB42495
 ID AAB42495 standard; Protein: 650 AA.

AC AAB42495;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2259 polypeptide sequence SEQ ID NO:4518.

XX Human: open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vlnery; antipsoptic; antiparkinsonian; noctropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antihypertensive;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antineoplastic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.

OS WO200058473-A2.

PN 05-OCT-2000.

PD 31-MAR-2000; 2000WO-US08621.

PE 31-MAR-1999; 9905-0127607.

PR 02-APR-1999; 9905-0127636.

PR 05-APR-1999; 9905-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI: 2000-602362/57.

XX N-PSDB; AAC76704.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 3717-3718; 5507pp; English.
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vlnery;
 CC antipsoptic; antiparkinsonian; noctropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiviral; antibacterial; antifungal; antirheumatic;
 CC antithyroid; and antineoplastic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

SO Sequence 650 AA;

Query Match 61.6%; Score 3229; DB 21; Length 650;
 Best Local Similarity 98.4%; Pred. No. 3.3e-289;
 Matches 633; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 397 ELTAFESATADASPFPAACKPVVYVSSLLLOEEPLAGKRGADGSGLEAVRLGPPSSGL 456
 DB 8 ELTAFESATADASPFPAACKPVVYVSSLLLOEEPLAGKRGAGGSGLEAVRLGPPSSGL 67
 QY 457 LVDMLEMDPVPVSSCPQLQRLFLFSRRKGGAQVPSFRVYLTLTFHQSSWPTLHOCI 516
 DB 68 LVDMLEMDPVPVSSCPQLQRLFLFSRRKGGAQVPSFRVYLTLTFHQSSWPTLHOCI 127
 QY 517 RVLLGKSREQRPDPASASIDFLMACIHVRPIWQGRDQTPQKRREELVLRVQGPPLISVE 576
 DB 128 RVLLGKSREQRPDPASASIDFLMACIHVRPIWQGRDQTPQKRREELVLRVQGPPLISVE 187
 QY 577 LILAEFRSDGDTAASCLIOARPLLLSCCGDSEVRVTEHLSCCICQWQDSVLR 636
 DB 188 LILAEFRSDGDTAASCLIOARPLLLSCCGDSEVRVTEHLSCCICQWQDSVLR 247
 QY 637 RCRLDLDLYLQRPRLRPVPEVLLHSEGAASSVCKIDGLHRTTLADTSRALEN 696
 DB 248 RCRLDLDLYLQRPRLRPVPEVLLHSEGAASSVCKIDGLHRTTLADTSRALEN 307
 QY 697 RGADASMCRKLAVAHPLLLRHLPMTAALLHGRTHLNFOEFRQONHLSCEFLVGLLEL 756
 DB 308 RGADASMCRKLAVAHPLLLRHLPMTAALLHGRTHLNFOEFRQONHLSCEFLVGLLEL 367
 QY 757 LQPHVFRSEHGALMDCLSTFRLILNKRKSSRHIAAFINFKVFIHXYTYNMPAALSF 816
 DB 368 LQPHVFRSEHGALMDCLSTFRLILNKRKSSRHIAAFINFKVFIHXYTYNMPAALSF 427
 QY 817 LQKHADPLHDSFDSNDVLMKSLLAGSLPSRDDRTRGDEEEESSAGSLPLVSVS 876
 DB 428 LQKHADPLHDSFDSNDVLMKSLLAGSLPSRDDRTRGDEEEESSAGSLPLVSVS 487
 QY 877 LFTPLTAEMAPYMKRISRGQIVE-----DLEVLSDIDMSRRPEIISFSTN 926
 DB 488 LFTPLTAEMAPYMKRISRGQIVEEGSPASPTDLEVLSDIDMSRRPEIISFSTN 547
 QY 927 LQRLMSSAECCRNIAFSLARSMONSPSIAAFLPTMYCLGSDDFVOTALRNLPY 986
 DB 548 LQRLMSSAECCRNIAFSLARSMONSPSIAAFLPTMYCLGSDDFVOTALRNLPY 607
 QY 987 ALLCOEHAVALHRAFLVGMGOMPASQISEALRIILHMEAVM 1029
 DB 608 ALLCOEHAVALHRAFLVGMGOMPASQISEALRIILHMEAVM 650

RESULT 5

AA87196
 ID AAY87196 standard; Protein: 404 AA.
 AC AAY87196;
 XX
 DT 09-MAY-2000 (first entry)
 XX
 DE Human secreted protein sequence SEQ ID NO:235.
 XX
 KM Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KM antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
 KM tumour; neurodegenerative disorder; developmental abnormality; allergy;
 KM foetal deficiency; blood disorder; immune system disorder; arthritis;
 KM autoimmune disease; hepatic disease; renal disease; inflammation;
 KM Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
 KM infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 KM asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 KM reproductive disorder; gastrointestinal disorder; respiratory disorder;
 KM metabolic disorder; food additive; preservative.
 XX
 OS Homo sapiens.
 XX
 PN WO200004140-A1.
 XX
 PD 27-JAN-2000.
 XX
 PF 14-JUL-1999; 99WO-US15849.
 XX
 PR 15-JUL-1998; 98US-0092921.
 PR 15-JUL-1998; 98US-0092922.
 PR 15-JUL-1998; 98US-0092956.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
 PI Mucenski M, Endress GA, Soppet DR;
 XX
 DR WPI: 2000-161128/14.
 DR N-PSDB: AAY98062.
 XX
 XX New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers, neurological or blood disorders -
 PT
 PS Disclosure; Page 466-467; 494pp: English.
 XX
 XX The polynucleotide sequences given in AAY98017 to AAY98108 encode the
 CC human secreted proteins given in AAY87064 to AAY87223. Human secreted
 CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Human secreted proteins and their polynucleotides can
 CC be used for developing products for the diagnosis or treatment of cancer,
 CC tumours, neurodegenerative disorders, developmental abnormalities and
 CC foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The
 CC proteins or polynucleotides can also be used as food additives or
 CC preservatives. The proteins are also useful for identifying their
 CC binding partners. AAY98008 to AAY98016 and AAY87063 are sequence used in
 CC the exemplification of the present invention.
 XX
 SO Sequence 404 AA;

Query Match 39.2%; Score 2055; DB 21; Length 404;
 Best Local Similarity 100.0%; Pred. No. 6,1e-181;
 Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 IQQWGSVIGRRCDLLDLYLQRPPELRVPPVLLHSEGAASSVCKLDGLHRTITLL 665
 DB 1 IQQWGSVIGRRCDLLDLYLQRPPELRVPPVLLHSEGAASSVCKLDGLHRTITLL 60
 QY 686 ADTSDSRALENRGADASMCRKLAVAHPLDLLLHPLMIALHGRHNLNFEFRQONHLS 745
 DB 61 ADTSDSRALENRGADASMCRKLAVAHPLDLLLHPLMIALHGRHNLNFEFRQONHLS 120
 QY 746 CPHLVGLLELLOPHYFRSEHOGALDCLISFTRLNLNKRKSRHAAFTNKVQTHY 805
 DB 121 CPHLVGLLELLOPHYFRSEHOGALDCLISFTRLNLNKRKSRHAAFTNKVQTHY 180
 QY 806 ITYNAPAAISFLQKHADPLHDSFSDSLVLMKSLAGSLSPRDRTPRGDEEEES 865
 DB 181 ITYNAPAAISFLQKHADPLHDSFSDSLVLMKSLAGSLSPRDRTPRGDEEEES 240
 QY 866 SAGSLPLVSVSLFTPLTAAEMAPYMKRLSGQTVEDLLEVLSDIDMSRRRPEILSFFST 925
 DB 241 SAGSLPLVSVSLFTPLTAAEMAPYMKRLSGQTVEDLLEVLSDIDMSRRRPEILSFFST 300
 QY 926 NLQRLMSAECCRNLAFLSLALRSMONSPSIAAFLPTFMYCIGSDFEVYQALNLEPE 985
 DB 301 NLQRLMSAECCRNLAFLSLALRSMONSPSIAAFLPTFMYCIGSDFEVYQALNLEPE 360
 QY 986 YALLCOEHAVALHRAFLVGMVGMPSAOISALRLHMEAVN 1029
 DB 361 YALLCOEHAVALHRAFLVGMVGMPSAOISALRLHMEAVN 404

RESULT 6
 AAE06173
 ID AAE06173 standard; Protein: 404 AA.
 AC AAE06173;
 DT 24-SEP-2001 (first entry)
 XX
 DE Human gene 46 encoded secreted protein fragment, SEQ ID NO:235.
 XX
 KM Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 KM foetal abnormality; developmental abnormality; haematopoietic disorder;
 KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KM Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KM psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KM inflammation; neurological disorder; Alzheimer's disease; food additive;
 KM angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KM pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KM cell culture; chemotaxis; vulnery; binding partner identification;
 XX
 OS Homo sapiens.
 XX
 PN WO200151504-A1.
 XX
 PD 19-JUL-2001.
 XX
 PF 12-JAN-2001; 2001WO-US00911.
 PF 13-JAN-2000; 2000US-0482273.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Mucenski M, Ebner R;
 XX
 DR WPI: 2001-425865/45.
 DR
 PT Isolated nucleic acid molecule encoding a human secreted protein is

PT used in preventing, treating or ameliorating a medical condition
XX
XX Disclosure: Page 116; 864pp; English.

CC AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
CC protein genes, and AAE06041-AAE06132 represent the proteins they encode.
CC AAE06133-AAE06205 represent human secreted protein fragments.
CC The secreted proteins and their genes are useful for preventing, treating
CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 71 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC The present sequence represents a human secreted protein fragment
CC referred to in the disclosure of the invention.
XX
XX

Sequence 404 AA:

Query Match 39.2%; Score 2055; DB 22; Length 404;
Best Local Similarity 100.0%; Pred. No. 6,1e-181;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 IQQMGSVIGRCRDLLDLYLQRPRLRVPEVLLHSGAASSYCKLDGLHRTTL 685
DB 1 IQQMGSVIGRCRDLLDLYLQRPRLRVPEVLLHSGAASSYCKLDGLHRTTL 60
QY 686 ADTSDBALENGGADASMAKRLAVAPLLLRHLMALHGRHLPNOEPRQNHLS 745
DB 61 ADTSDBALENGGADASMAKRLAVAPLLLRHLMALHGRHLPNOEPRQNHLS 120
QY 746 CFLHVLGELLELPYHFRSEHOGALMDCLSFIRLLNTRKSSRHIAAEINKFVQFIHRY 805
DB 121 CFLHVLGELLELPYHFRSEHOGALMDCLSFIRLLNTRKSSRHIAAEINKFVQFIHRY 180
QY 806 IYVNPAAISFLOKHAADPLHDSFNSDIYMLKSLAGSLPSRDRTRGDEEEES 865
DB 181 IYVNPAAISFLOKHAADPLHDSFNSDIYMLKSLAGSLPSRDRTRGDEEEES 240
QY 866 SAGSLPIVSVSFTPLTAEMAPYMKRLSRGQVEDLLLEVLSPIDMSRRRPETLSFFST 925
DB 241 SAGSLPIVSVSFTPLTAEMAPYMKRLSRGQVEDLLLEVLSPIDMSRRRPETLSFFST 300
QY 926 NIQRLMSSAECCRNLAESLARSMONSPSIAAFLPTMYCIGSODFEVOTALRNPDE 985
DB 301 NIQRLMSSAECCRNLAESLARSMONSPSIAAFLPTMYCIGSODFEVOTALRNPDE 360
QY 986 YALLCOEHAAYVLLHRAFLVGMGMDPSAOISEALRIIEMEAVM 1029
DB 361 YALLCOEHAAYVLLHRAFLVGMGMDPSAOISEALRIIEMEAVM 404

RESULT 7
ABG33997
ID ABG33997 standard; Protein; 404 AA.
XX

AC ABG33997;
XX
XX 15-JUL-2002 (first entry)

DE Human secreted protein #2 encoded by gene 46.

KW Human: secreted protein; gene therapy; immunosuppressive;
XX antiarthritis; antirheumatic; antiproliferative; cytostatic; cardiac;
XX vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
KW viruticide; fungicide; opthalmological; autoimmune disease; neoplasm;
KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; skin aging;
KW epithelial cell proliferation; food additive.

OS Homo sapiens.

PN MO200226931-A2.

PD 04-APR-2002.

PF 24-SEP-2001; 2001WO-US29871.

PR 25-SEP-2000; 2000US-234925P.

PR 12-JAN-2001; 2001WO-US00911.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;
PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
PI Endress GA, Mucenski M, Eder R;

PI WPI: 2002-362489/39.

DR N-PSDB: ABR69771.

PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
PT the polypeptides, useful for treating Huntington's disease, sepsis,
PT meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
PT asthma

XX Disclosure: Page 1313-1314; 1478pp; English.

XX The invention relates to an isolated nucleic acid molecule (or its
XX fragment, homologue complement or allelic variant) encoding a human
XX secreted protein (and its fragment, domain, epitope, variant, secreted
XX form and species variant). Also included are a recombinant vector
XX comprising the nucleic acid, a recombinant host cell comprising the
XX vector, an antibody against the secreted protein, a recombinant host cell
XX that expresses the secreted protein and a method of identifying a binding
XX partner of the secreted protein. The nucleic acid and protein are used to
XX prevent, diagnose, treat or ameliorate a medical condition in e.g.
XX humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
XX for example autoimmune diseases e.g. rheumatoid arthritis,
XX hyperproliferative disorders e.g. neoplasms of the breast, or liver,
XX cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
XX e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
XX Alzheimer's disease, infections caused by bacteria, viruses and fungi and
XX ocular disorders e.g. corneal infection. Many other diseases and
XX disorders are listed in the specification. The polypeptides can also be
XX used to aid wound healing in epithelial cell proliferation, to prevent
XX skin aging due to sunburn, to maintain organs before transplantation, for
XX supporting cell culture of primary tissues, to regenerate tissues and in
XX chemotaxis. The polypeptides can also be used as a food additive or
XX preservative to increase or decrease storage capabilities. The present
XX sequence represents a novel human secreted protein of the invention.

Query Match 39.2%; Score 2055; DB 23; Length 404;
Best Local Similarity 100.0%; Pred. No. 6,1e-181;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 IQQWGSVLRGRCDLLDQLYLRPELRVPEVYLHSEGAASSVCKDGLHRTFTLL 685
 DB 1 IQQWGSVLRGRCDLLDQLYLRPELRVPEVYLHSEGAASSVCKDGLHRTFTLL 60
 QY 686 ADSDSALNRGADAMACRKLAVAPHLRLPMLALHGRTHLNQEPFQOQNHLS 745
 DB 61 ADSDSALNRGADAMACRKLAVAPHLRLPMLALHGRTHLNQEPFQOQNHLS 120
 QY 746 CFLHVLGLLELLOPHVRSEHOGALMDCLLSFIRLLNRYKRSRHLAFTINKFVQFIHKY 805
 DB 121 CFLHVLGLLELLOPHVRSEHOGALMDCLLSFIRLLNRYKRSRHLAFTINKFVQFIHKY 180
 QY 806 ITYNAPAAISFLQKHAPRLDLSFNSDLYMKSLLAGLSLPSADDTDRGLDEGEES 865
 DB 181 ITYNAPAAISFLQKHAPRLDLSFNSDLYMKSLLAGLSLPSADDTDRGLDEGEES 240
 QY 866 SAGSLPVSVSLFPTLAAEMAPYMKRLSGQVYEDLEVLSDIDEMSRRPETLSFST 925
 DB 241 SAGSLPVSVSLFPTLAAEMAPYMKRLSGQVYEDLEVLSDIDEMSRRPETLSFST 300
 QY 926 NLQRLMSAECCRNLAFLSLARSMONSPSTIAAFLEPTFMVCLGSDPEVVOALRNLP 985
 DB 301 NLQRLMSAECCRNLAFLSLARSMONSPSTIAAFLEPTFMVCLGSDPEVVOALRNLP 360
 QY 986 YALLCOHAVALHRAFLVNGMGMDPSAOISEALRTIHMEAVM 1029
 DB 361 YALLCOHAVALHRAFLVNGMGMDPSAOISEALRTIHMEAVM 404

RESULT 8

AAE06086 standard; Protein: 327 AA.

24-SEP-2001 (first entry)

Human gene 46 encoded secreted protein HUBC135, SEQ ID NO:148.

Human: secreted protein; proliferative disorder; cancer; tumour; asthma;
 foetal abnormality; developmental abnormality; haematopoietic disorder;
 immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 inflammation; neurological disorder; Alzheimer's disease; food additive;
 angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 pregnancy-related disorder; endocrine disorder; infection; wound healing;
 cell culture; chemotaxis; vulnerrary; binding partner identification;
 gene therapy.

Homo sapiens.

Location/Qualifiers

Peptide 1..23 /label= Signal_peptide

Protein 24..327 /label= Mature_human_secreted_protein

W0200151504-A1.

19-JUL-2001.

12-JAN-2001; 2001WO-US00911.

13-JAN-2000; 2000US-0482273.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y,
 Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 Endress GA, Muscenski M, Ebner R;

WPI: 2001-425865/45.

DR N-PSDB; AAD11675.
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 used in preventing, treating or ameliorating a medical condition -
 Claim 11, Page 771-772; 864pp; English.

AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
 protein genes, and AAE06041-AAE06132 represent the proteins they encode.
 AAE06133-AAE06205 represent human secreted protein fragments.
 The secreted proteins and their genes are useful for preventing, treating
 or ameliorating medical conditions, e.g., by protein or gene therapy.
 Pathological conditions can be diagnosed by determining the amount of the
 new protein in a sample or by determining the presence of mutations in
 the new genes. Specific uses are described for each of the 71 genes,
 based on the tissues in which they are most highly expressed, and include
 developing products for the diagnosis or treatment of proliferative
 disorders, cancer, tumours, foetal and developmental abnormalities,
 haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 angiogenic disorders, kidney disorders, gastrointestinal disorders,
 pregnancy-related disorders, endocrine disorders, and infections. The
 proteins can also be used to aid wound healing and epithelial cell
 proliferation, to prevent skin aging due to sunburn, to maintain organs
 before transplantation, for supporting cell culture of primary tissues,
 to regenerate tissues, to identify their cognate ligands or binding
 partners, and in chemotaxis, and can be used as a food additive or
 preservative to modify storage properties. Antibodies specific for a
 protein of the invention can be used in alleviating symptoms associated
 with the disorders mentioned above, and in diagnostic immunoassays e.g.,
 radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
 The present sequence represents a human secreted protein of
 the invention.

Sequence 327 AA;

Query Match 31.8%; Score 1665; DB 22; Length 327;

Best Local Similarity 100.0%; Pred. No. 5.2e-145;

Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 703 MACRKLAVAPHLRLPMLALHGRTHLNQEPFQOQNHLSCELVGLLELLOPHV 762

DB 1 MACRKLAVAPHLRLPMLALHGRTHLNQEPFQOQNHLSCELVGLLELLOPHV 60

QY 763 RSEHOGALMDCLLSFIRLLNRYKRSRHLAFTINKFVQFIHKYTYNAPAAISFLQKHAD 822

DB 61 RSEHOGALMDCLLSFIRLLNRYKRSRHLAFTINKFVQFIHKYTYNAPAAISFLQKHAD 120

QY 823 PLHDLSDNSDLYMKSLLAGLSLPSRDRDTRDGLDEGEESAGSLPVSVSLFPTLT 882

DB 121 PLHDLSDNSDLYMKSLLAGLSLPSRDRDTRDGLDEGEESAGSLPVSVSLFPTLT 180

QY 883 AAEMAPYMKRLSGQVYEDLEVLSDIDEMSRRPETLSFSTNLQRLMSAECCRNLA 942

DB 181 AAEMAPYMKRLSGQVYEDLEVLSDIDEMSRRPETLSFSTNLQRLMSAECCRNLA 240

QY 943 FSLARSMONSPSTIAAFLEPTFMVCLGSDPEVVOALRNLPALCOHAVALHRAFL 1002

DB 241 FSLARSMONSPSTIAAFLEPTFMVCLGSDPEVVOALRNLPALCOHAVALHRAFL 300

QY 1003 LVGMYGMDPSAOISEALRTIHMEAVM 1029

DB 301 LVGMYGMDPSAOISEALRTIHMEAVM 327

RESULT 9
 AAG33908
 ID AAG33908 standard; Protein: 327 AA.

AC AAG33908;

XX 15-JUL-2002 (first entry)
 XX
 DE Human secreted protein encode by gene 46.
 KW Human; secreted protein; gene therapy; immunosuppressive;
 KW antithetic; antirheumatic; antiproliferative; cytostatic; cardiant;
 KW vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
 KW virucide; fungicide; ophthalmological; autoimmune disease; neoplasm;
 KW rheumatoid arthritis; hyperproliferative disorder; cerebral arrest;
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing; skin aging;
 KW epithelial cell proliferation; food additive.
 XX
 OS Homo sapiens.
 PN WO200226931-A2.
 PD 04-APR-2002.
 PF 24-SEP-2001; 2001WO-US29871.
 PR 25-SEP-2000; 2000US-234925P.
 PR 12-JAN-2001; 2001WO-US00911.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Mucenski M, Ebner R;
 XX
 DR N-PSDB; ABR69771.
 XX
 PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
 PT the polypeptides, useful for treating Huntington's disease, sepsis,
 PT meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
 PT asthma
 XX
 PS Claim 11: Page 1250-1251; 1478pp: English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (or its
 CC fragment, homologue complement or allelic variant) encoding a human
 CC secreted protein (and its fragment, domain, epitope, variant, secreted
 CC form and species variant). Also included are a recombinant vector
 CC comprising the nucleic acid, a recombinant host cell comprising the
 CC vector, an antibody against the secreted protein, a recombinant host cell
 CC that expresses the secreted protein and a method of identifying a binding
 CC partner of the secreted protein. The nucleic acid and protein are used to
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
 CC for example autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
 CC ocular disorders e.g. corneal infection. Many other diseases and
 CC disorders are listed in the specification. The polypeptides can also be
 CC used to aid wound healing an epithelial cell proliferation, to prevent
 CC skin aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities. The present
 CC sequence represents a novel human secreted protein of the invention.
 XX
 SQ Sequence 327 AA;
 Query Match 31.8%; Score 1665; DB 23; Length 327;
 Best Local Similarity 100.0%; Pred. No. 5, 2e-145;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 703 MACRKLAVAHPLLLRLHPLMTAALLHGRTHLNFGDFRQONHLSCLFVLHGLLELLQPHVF 762

DB 1 MACRKLAVAHPLLLRLHPLMTAALLHGRTHLNFGDFRQONHLSCLFVLHGLLELLQPHVF 60
 QY 763 RSEHOGALMDCLLSFIRLLNWRKSSRHLLAFLINKFVFIKITYTYNPAAPISFLQKAD 822
 DB 61 RSEHOGALMDCLLSFIRLLNWRKSSRHLLAFLINKFVFIKITYTYNPAAPISFLQKAD 120
 QY 823 PLHDSFNSDLVWMLKSLAGLSLPSRDRDRDGLDEGEESAGSLPLVSVSLFPLT 882
 DB 121 PLHDSFNSDLVWMLKSLAGLSLPSRDRDRDGLDEGEESAGSLPLVSVSLFPLT 180
 QY 883 AAEMAPYMKRLSRQTVEDLLEVLSDIDEMSRRREILSFSTNLQRLMSSAECCRNLA 942
 DB 181 AAEMAPYMKRLSRQTVEDLLEVLSDIDEMSRRREILSFSTNLQRLMSSAECCRNLA 240
 QY 943 FSLALRSMONSPSTAAALPTFMVCLGSDQDFEVQATARNLPEVALLCQEHAAVILHRAE 1002
 DB 241 FSLALRSMONSPSTAAALPTFMVCLGSDQDFEVQATARNLPEVALLCQEHAAVILHRAE 300
 QY 1003 LVGMVGMDFPSAQISEALRIHMEAYM 1029
 DB 301 LVGMVGMDFPSAQISEALRIHMEAYM 327
 RESULT 10
 AAY87109
 ID AAY87109 standard; Protein: 328 AA.
 AC AAY87109;
 DT 09-MAY-2000 (first entry)
 XX
 DE Human secreted protein sequence SEQ ID NO:148.
 XX
 KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antinflammatory; nootropic; neuroprotective; antiallergic; cancer;
 KW tumour; neurodegenerative disorder; developmental abnormality; allergy;
 KW foetal deficiency; blood disorder; immune system disorder; arthritis;
 KW autoimmune disease; hepatic disease; renal disease; inflammation;
 KW Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
 KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 KW asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 KW reproductive disorder; gastrointestinal disorder; respiratory disorder;
 KW metabolic disorder; food additive; preservative.
 XX
 OS Homo sapiens.
 XX
 PN WO200004140-A1.
 PD 27-JAN-2000.
 PF 14-JUL-1999; 99WO-US15849.
 PR 15-JUL-1998; 98US-0092921.
 PR 15-JUL-1998; 98US-0092922.
 PR 15-JUL-1998; 98US-0092956.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
 PI Mucenski M, Endress GA, Soppet DR;
 XX
 DR WPI: 2000-161128/14.
 DR N-PSDB: AA298062.
 XX
 PT New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers, neurological or blood disorders
 XX
 PS Claim 11: Page 404-406; 494pp: English.
 CC The polynucleotide sequences given in AAY86017 to AAY898108 encode the
 CC human secreted proteins given in AAY87064 to AAY87223. Human secreted

CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Human secreted proteins and their polynucleotides can
 CC be used for developing products for the diagnosis or treatment of cancer,
 CC tumours, neurodegenerative disorders, developmental abnormalities and
 CC foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The
 CC proteins or polynucleotides can also be used as food additives or
 CC preservatives. The proteins are also useful for identifying their
 CC binding partners. AA298008 to AA298016 and AA87063 are sequence used in
 CC the exemplification of the present invention.

CC XX Sequence 328 AA:

Query Match 31.8%; Score 1665; DB 21; Length 328;
 Best Local Similarity 100.0%; Pred. No. 5.2e-145;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 703 MACRKAIVAHPLILRLHRLMIAALHGRTHLNQFROONHLSCLFVHVLGILLLOPHVF 762
 DB 1 MACRKAIVAHPLILRLHRLMIAALHGRTHLNQFROONHLSCLFVHVLGILLLOPHVF 60
 QY 763 RSEHOGALWDCLLSFIRLLLNKRSRHLAFLNKFVQFIHKYITYNAPAAISFLQKHD 822
 DB 61 RSEHOGALWDCLLSFIRLLLNKRSRHLAFLNKFVQFIHKYITYNAPAAISFLQKHD 120
 QY 823 PLHDLSDVLMKSLAGLSLPSRDRDTBGDLDEGEESASGLPLVSVSLTPPLT 862
 DB 121 PLHDLSDVLMKSLAGLSLPSRDRDTBGDLDEGEESASGLPLVSVSLTPPLT 180
 QY 883 AAEAPMYMRLSRGQVEDLEVLSDIDEMSRRLPILSFESNTLRLSSAECCRNLA 942
 DB 181 AAEAPMYMRLSRGQVEDLEVLSDIDEMSRRLPILSFESNTLRLSSAECCRNLA 240
 QY 943 FSLALRSMONSPSIAAFLPTEMYCLGSQDFEVVQALRNLPEYALLQGEHAAVLLHRAF 1002
 DB 241 FSLALRSMONSPSIAAFLPTEMYCLGSQDFEVVQALRNLPEYALLQGEHAAVLLHRAF 300
 QY 1003 LVGMVGOMPSAOISFALRIHMEAVM 1029
 DB 301 LVGMVGOMPSAOISFALRIHMEAVM 327

RESULT 11
 ABB59689
 ID ABB59689 standard; protein; 2042 AA.

XX ABB59689;
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 5859.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 OS Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX

PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 XX N-PSDB; ABL03792.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 5859; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175) and the encoded proteins
 CC sequences (AB01840-AB01840) and the encoded proteins
 CC (AB01840-AB01840).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIP
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC XX Sequence 2042 AA:

Query Match 16.5%; Score 863; DB 22; Length 2042;
 Best Local Similarity 25.4%; Pred. No. 1.5e-69;
 Matches 280; Conservative 191; Mismatches 402; Indels 230; Gaps 32;

QY 1 MHLVYVHAWYILTLGPRRDSFQALLDIWPEEKPLPTAFLVOTSEAL-LTDDWK 59
 DB 1084 LHLNIIHAFILTLTYS--NSMPESTIPILDYPPRPAPAVAPSPDQVOLLDPWK 1141
 QY 60 LRMIRESVLRVDAALDLEPQDLLEFVSGFIPVSMKTLQFLDQAVAHADQTEQNT 119
 DB 1142 LKMISSVRLIEALNDLTPDQIVLFVQNGFVMSKLLAMDPAVEQDVLKMI 1201
 QY 120 MOKNYAHLEVOHERGASGOTFHSLLTASLPPRRDSTEARPKSSPEQIPQGRIRNG 179
 DB 1202 LNKAYLAQLLEIQARGANG--HYVOALDHSQTPDLPKTS--VVIOEAVEI- 1254
 QY 180 TQRLVGPEDDLAQMFLQIFPLSPDRMOSSPRPAALQALQGLARVVGSGSEVPG 239
 DB 1255 -----DDYDS-----SDSDRPPTNPLATKKEVAQTILTPOQLTESNS 1291
 QY 240 ITFVRLQALATLTLSPH-----GALVMSMRSHFLACPLRLQOLCOYOR 283
 DB 1292 DCRSLIQLKLLDMLASNSNRADVNAITTEVLAVGCVTMSRH-----ACFELR----- 1339
 QY 284 CVPQDTGFSSLEFLKVLQMLQDLSDPGVEGGLPRAQLRLASQASGRRLSDVRCGLRL 343
 DB 1340 -----TFEFCMLHSDKYHILE-----NALQKMLSMF-----KTFADSSILQK 1377
 QY 344 AELALFRODLEVVSVRAVIATLRSGECCSVPELDS--KVLQGLEVNSPLLEELTA 401
 DB 1378 SE--LYHESLVFMLRNSREIYA-----QOFKANTALVARKRIVRAIVO----- 1418
 QY 402 FFSATADAASPPACKPVVVVSSILLQEEPLAGKPGADGGSLEAVRLGSSGLLVDM 461
 DB 1419 SFDQTKDS-----KTVAKSKSDQLFH-----NGFLPDM 1447
 QY 462 EMIDPEVSSCPDLQL---RLFSRRKGQQAQVPSFRPLTLTFTHQSSWPLHOCIRV 518
 DB 1448 SEMDPEIVSR---QLMKERFLFSK-----SCSEFRYLLSLINHQNMWDIEI-IAE 1495
 QY 519 LLGKSREGRDPSASLDPLMACIHVRIMGROGRTPQKRRELTVLRVGPDELISLVELI 578

Db 1496 YLFKHFHEDYATVATVLANFEALTTNPKLKGRKYMKNVRBDAPFMRITSELPESHFI 1555
QY 579 ---LAEATRSODGDTAGSLIQARPLLLSCCGDDESVKAYVEHLSGCGQAGDSVL 634
Db 1556 LHEGISVEKLDKSNDFELCS---RMNLEFKLEKRRDLWKVEHVE-----KSSVS 1605
QY 635 GRRCDDLQQLYLRPELRVPEVLLHSEGAASSV-----CKLDGLHREFTLLA 686
Db 1606 DYLAQVLAQVIMYPRIKFLKP-----GKTGQAKKLQKCKQADKVSNNLITCIG 1658
QY 687 DTSRSLAENMGADASMACRIAAVAPLLLRHLPMAALLHGRTHLNFQEFQGNHSC 746
Db 1659 SLVGKKDETLSTDTELLRLKIASHPLLFLROLGVLSIMOGRAQLSMKALREEHFHR 1718
QY 747 FLHVGLLELLQPHVF-----RSEHGALMDCLISFIRLLNKRKSSRHAAFINKEVQFI 802
Db 1719 FVOILRTLELLQPTIFEEAYKNEIONTL-SCYFNFK---HNSVKEACQMLNKFVQWL 1773
QY 803 HKYITYNAPPAISFLQKHADPLHDLSPDNSDLYMLKSLAGLSLPSRDDRTDGLDEEG- 861
Db 1774 QAYINYNSSALLFTEQVVGILKELAAVYTSIGKLOVLAQVALLQHKSHSATFELDEY 1833
QY 862 -----EESASGSLPVSYSLFTPLIAEMA 887
Db 1834 KYEYDLDEHFPVKSPKPVYTEDEPIEVNPTPIIDPSSRGPISVLTIGYSRSNYTDS 1893
QY 888 PY---MKRLSRGOTVEDLLEVLSDIDMSRRRPETLSFSTINLQRLSSAECCRNIAF 943
Db 1894 PHFLDLVATIKOSNTEDEVVLGPMOELCTLSKRFVFINLEFRLNLTFPSQAIRSTAF 1953
QY 944 SLAISMQNSP---SIAAFLPTFMYCIGSODEFVYQALRNLPYALLCQEHAAVLLHR 1000
Db 1954 IILRHLEKHNFGNSDINICTINAVIQCLRDENSSVAATAIDNLPMSVLDJHAIDILTV 2013
QY 1001 AFLVGMGOMPSSQISALRL 1023
Db 2014 AFSIGLKSCLNTGHQIRKVLQTL 2036

RESULT 12
AAM95497
ID AAM95497 standard; Protein: 161 AA.
XX
AC AAM95497;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen SEQ ID NO: 4155.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
XX
KW cancer; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234222.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0239678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI: 2001-465570/50.
 DR N-PSDB: AAL01467.
 XX
 XX Isolated nucleic acid molecule encoding a reproductive system antigen
 PT is used in preventing, treating or ameliorating a medical condition -
 XX
 XX
 PS Claim 11: SEQ ID NO 4155; 1297pp + Sequence Listing: English.
 XX
 XX The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a protein of the invention.
 CC
 XX Sequence 161 AA:
 SQ

Query Match 10.1%; Score 530.5; DB 22; Length 161;
 Best Local Similarity 73.1%; Pred. No. 1.5e-40;

Matches 117; Conservative 7; Mismatches 17; Indels 19; Gaps 3;
 QY 812 AATSFLOKADPLHDSFSDSLVMLKSLAGSLPRDRTDGLDEDEGEESAGSLP 871
 |||
 Db 2 AATSFLOKADPLHDSFSDSLVMLKSLAGSLPRDRTDGLDEDEGEESAGSLP 61
 QY 872 LVSVLEFPLTAAEMAPYMKRLSGQVE-----DLLEVLSDIDEMSRRPETLS 921
 |||
 Db 62 LVSVLEFPLTAAEMAPYMKRLSGQVEGESXPXSTPDLLESLDIDEMSRRPETLS 121
 QY 922 FFSTNLQRLMSSAE-----CCRNLAFLRLRMQNSPSI 956
 |||
 Db 122 FFEDQ---PAADELPRXXCPXPPLXPALRSCRTAPAL 157
 RESULT 13
 ID AAU33568 standard; Protein; 2472 AA.
 XX
 AC AAU33568;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Pseudomonas aeruginosa cellular proliferation protein #12.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001MO-US09180.
 PF
 XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR N-PSDB: AAS51427.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 PT
 XX
 PS Example 3; Seq ID NO 5064; 51pp; English.
 XX
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic
 CC format directly from WIFO at
 CC ftp.wifo.int/pub/published_pct_sequences.
 XX

SO Sequence 2472 AA;

Query Match 3.0%; Score 156; DB 22; Length 2472;
 Best Local Similarity 19.1%; Pred. No. 0.00052;
 Matches 229; Conservative 142; Mismatches 386; Indels 440; Gaps 53;

```

QY 8 AMVILLITGPPRADSEFOALDIDWPEEKPLTAFVDTSE-EAL--LIPDM----- 57
DB 846 AONVSVSLPPADE-----PVDEELREVEIEEAGEVLETRIGRYLPAKMDHDD 895
QY 58 -----LK--LRMRSEVL-----RVYVAALQDLERPOQLLFVQSF 90
DB 896 REALTEVRRATHTLKGSGMVRALVIGELAWSIENLFNNHVRDSTIASPEVO----- 947
QY 91 GIPVSSMSKLLQFLDQAAVANDPQTEONIMDKNMAHLVEVOHERGASGGQTFHSILTAS 150
DB 948 -----RVVDQVVALPELVEE-----FAANAQRDRDVIDLLAATAHALAKE 989
QY 151 ---LPPRDSTEARPKSSPEOPIC-----QGRIRVGTQLRVLCPEDDIAG 193
DB 990 PLPEPPADDGQVP--PEAGAEPSSLDNGVQAPLADAPQAAMAEQSDVELLDPO----- 1043
QY 194 MFLQFFPLSPDRMS-----SSPRPALAQOLGELARVQSGSEVVGITVR 243
DB 1044 -LLEFTNEAETHLALVGLADCARLPQVTDALQRL-----HTLKGSAHMGIL-- 1095
QY 244 VLQALATLSSPGALVMSHRSHFLA-----CPLRLQCYQRCVPQDTFFSSILFKV 298
DB 1096 PLAEIATPLEK-----LVKEYKSNLAFDLAEALHDAEQLFRGLEQVQAGRLPI 1149
QY 299 -----LLOMLQMDSPGV-----EGGPLRAQIRMLASQASAGRSLDVRGGLRIA 344
DB 1150 PGSDALLERIEALHOERIASLEAERYSDAGERRDPLLEAFVLEGMDILLDAEDLIERWH 1209
QY 345 EALAFRODLEVVSSTVRAVIATLRSG-----EGGSVPRDLISKVGLLI----- 388
DB 1210 EHPQERQEL---SALREELSTLDRGARHAELPQVEELQALALLADVAEEERLAVSPAF 1265
QY 389 --EVSRPH-----LEELLTA-----FFSATADAASPP 414
DB 1266 FEARQAHAHALGMMDOVAGIQTTPRPERVALQELLEAPAAEAVPIFDPSLADDDP 1325
QY 415 -----ACKRVVVVSSLLQEEEPPLAGKPGADGSL-----EAVRLGPSSG-LIV 458
DB 1326 PEDEEPALPEAVYEAGAPAEETVPAAPAPAGRELEDEEMVSIFFLEAVDILESAGQALA 1385
QY 459 DWLEMLDPEVSSCPDLOLRLEFSRRKKGQAOVPSFRYLLTLFTHOSWFTLHCIRIV 518
DB 1386 QW--QAEFGAUSLSALO-RDHLTLKGARMAELAEIGD-----LAHE-----LEA 1428
QY 519 LIGKSRQEFDPASLIDFLMACIHVPRIWQGRDQT---PQRREELVLRV---QGP-- 569
DB 1429 LYEGLVDRRYQHSPLAGLQACH-DRLAEQLDQLSAGAPLADPHLLIOSIRFRGPAVA 1487
QY 570 -----ELISVELLAEETRSQODGTACSLIQARLPLLISCCGDEDSYKRYTEHS 623
DB 1488 EAATFGAESPVLEIYAPA---VEEPAPAAAEFERDELEIFL---EEGFDIIDLSAA 1541
QY 624 GCIOQMGSV----- 633
DB 1542 AALQRMMDVDTTELEALQRLHLTKGARMATGEIGDLAHELEFLEGCGRLRAS 1601
QY 634 -----LGRRCRDLLQLY-----LGR-----PELRVVPPEVL---LH 662
DB 1602 PALFGILQCHDELAEMLAEVAGHRTLPDQGLALAEIRLRASDPQGLVSPVSXLKPLA 1661
QY 663 SGAASSSVCKLDGLIHRITTLAD-----TSRRLLENFGADASNAACKRLAVA 711
DB 1663 AGGAADSESIIDFLAEADDLLENLELALGRWDGNGDAQPLDD----- 1706

```

```

QY 712 HPLLRLHPMIALHGRTHLNFQEFROQNHLSCEFLHVLGLLELQHPVERSEOGALW 771
DB 1707 --LRLRLHLTKGARLAGTEIG-----NLAHLEQHLTLTAQOOGAPW 1747
QY 772 -DCLLSFIRLLNLYKSSRHIAAFINKFVOFTHKYITYNAPPAISFLQKHADPLHDSFD 830
DB 1748 PDSL-----LDAQGLBGLQROVDLLERLAE 1775
QY 831 NSDLYMKSLLAG-----LSLPSRDDTRD---GLDEGESESSAGSLPVYSLSLTP 880
DB 1776 DDE-----AGEPEPAQALVQADD-DRRAVASALAEITRLAPAGA--TMAEAPAP 1824
QY 881 LTAEMAPYMKRLSRQQTVEDELEVLSIDEMSRRP-EILSFFSTNQLRMSAAE 936
DB 1825 AAPATTLFPVKRAQEAQ-----EASRRAPQELVAVPALLEENLVNAGE 1870

```

RESULT 14
 AAM52358
 ID AAM52358 standard; Protein; 1217 AA.

AC AAM52358;

DT 25-JAN-2002 (first entry)

DE Putative TBP interactive protein.

KW Gemlinivirus; plant; viral infection; transgenic plant;

KW tomato yellow leaf curl virus.

OS Arabidopsis thaliana.

PN FR2806095-A1.

PD 14-SEP-2001.

PF 10-MAR-2000; 2000FR-0003140.

PR 10-MAR-2000; 2000FR-0003140.

PA (GENT-) GENTECH SARL.

PI Bejarano ER, Castillo GA, Colinet D, Donoso CI, Iniesta JR;

PI Grevesse C, Hericourt F;

DR WPI: 2001-628275/73.

DR N-PSDB: ABA01228.

PT New polynucleotides for producing transgenic plants resistant to
 PT geminivirus infection comprising polynucleotides encoding proteins
 PT which interact with at least one of the products of the geminivirus
 PT genome

PS Claim 4; Pages 44-48; 106pp; French.

CC The present invention relates to coding sequences encoding proteins which
 CC interact with at least one of the six products of the geminivirus genome
 CC necessary for infection of a plant by the virus. The present sequence is
 CC one such protein. The coding sequences are useful for producing
 CC transgenic plants resistant to geminivirus infection, particularly tomato
 CC yellow leaf curl virus.

SO Sequence 1217 AA;

Query Match 2.8%; Score 149; DB 22; Length 1217;
 Best Local Similarity 20.4%; Pred. No. 0.00074;
 Matches 196; Conservative 137; Mismatches 334; Indels 292; Gaps 49;

```

QY 46 DTSEALILPDMILKLMISEVRLVDAALQDLPEQQLLFPQSGEIPVSSMSKLLQFLD 105
DB 409 DTDESS---PKWL---LKQEVSKIYKSIINQLREKSVKTKGAFSV---LRELVLVLP 457

```

```

QY 106 QAVADPOTL-----EONINDKNTMAHL-VEVQERGASGQOTHSITLASLPPRRODTEA 160
DB 458 DCLADHIGSLVPGIERALANKSTSNKLE-----ALVFTKLVLASHAP----- 501
QY 161 P-----KPKSSPE-QPIGGGRIRV-----GTOLRVIGP----- 187
DB 502 PVFHPYIKALSSVLAAGVGRYKVTAEALRVGELRVVRPSTAGNGDFKPPVHETYN 561
QY 188 -----EDDLAGMFLQIFPLSPPRMOSSSPRPVALALQALGQELARVY 231
DB 562 AIMSRLTNDQDQVEKCAITCMGLVISTGDDLRALPSCLEPV---LVDRMGNEIRLT 618
QY 232 QGSEPEVPGITVRVQLATILSSPHGALVMSHRSHFACPLLRQLCYQRCVPODPTG 291
DB 619 -----AVKAFSVYATSP-----LH-----INLSCVLDHLIAELTGF 649
QY 292 SSLEKVLQMLQMLDPSGVEGGPRLRAQLMLASQASA-GRR-L-SDVRGGLRLAEALF 349
DB 650 LRKANRVLR-----ATLITMTLVTAIGDKIGSEAYEVILVELSLIS 693
QY 350 RODLEVSVSTRAVIAVLNRSGECSVEPDLI--SKVL-QGLIEVRSFHEEL---LTAF 402
DB 694 VSDLHMALALE-LCCTLMGKSCSENI SLAVRNKVLPAQLTLVKSPLLGQALLDQKF 752
QY 403 FSATA-DAAPFPACKRVVVVSSIL-LQEEPILAGKP-----GADCG 443
DB 753 FEALVYHANTSF-----YTLLESLSCKAPSPGSGVPKQALYSIACVAVLCIAGDN 807
QY 444 SLAVARLGPSSGLLVLMLEMDPEVWSSCPDLQRLRL---FSRRKG-KCOAQVPSRPY 498
DB 808 CSSTVAK-----LMELTKDSDGNSAKQHLALSLDEIGRRKLSHAGET----- 854
QY 499 LTLTTHQSSWPLTHOCITRVLCKSRQRPDSASLDFLMACIHVPRIMOGRDQRTPOKR 558
DB 855 -IYIESFQSPFEIKSNAASYALGNINAVNL--SNYLPEFI-----LDQIDMOCK 899
QY 559 REELVRYVQPELISLVELLAEAFRPSODGDNAACSLQARLPLILSCCGGDESRYK 618
DB 900 KQYTLHL-----SLKEVIVRO---SVDKADFQNSVETKLTALFNHCESEEGVRNV 948
QY 619 TEHLSCGIQMGDSVYLRRCRDLLQLYLQRPETLRVVPVLLHSEGAASSVC----- 672
DB 949 VARCLG-----KWLLEPEKLVPAALQVRTTSPAFAFRAVTVAVKY 989
QY 673 -----KLDGL-----IHRITTLADTSDSRALENRGADSMACRKLAVAH--PLILLRH 719
DB 990 SYVEREKDEIETFPQISSFLMLIKD-GDRHV--RRAAVALS---TFAHYKFNLIKGL 1042
QY 720 LPMIALLLHGRTHLNPEFQONHLSCFLHVL--GILELQPHVFRSEHOGALMDCLISF 777
DB 1043 LPELLEPLLDQTYIK-KELIRTYDLQPFKHVDDGL-----ELKRAAECEVFTL 1090
QY 778 IRLLLNRKSSRHAAFIINKFVQ-----FIHKYITYNAPAAI-SFLOKHADPLH- 825
DB 1091 VDSCLQDVNNSFIVPLKSGLEDHYDKMLCHILISLADKCSAVLAVLDSLVEPLHK 1150
QY 826 DLSFD-----NSDLVMLKSLLAGLSLPSRDRTDGLDEGEESSSAGSPL 872
DB 1151 TISFKPKQADAVKQEHDRNED--MIRSAIRAISSLDIRINGVDYSHKFGKMGDMKRSVPL 1207

```

```

RESULT 15
AA185660
ID AA185660 standard; Protein; 1519 AA.
XX
AC AA185660;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human tyrosine kinase substrate tks107/GRUB protein sequence.
XX
KM Tyrosine kinase substrate: Tks; Tks 107; GRUB: Tks 113; Tks 118; Dresh;
KM Tks 202; Src; Rheumatoid arthritis; atherosclerosis; stroke; cancer;

```

```

KW autoimmune disorder; organ transplantation; myocardial infarction;
KW cardiomyopathy; renal failure; neurodegenerative disorder;
KW cardiovascular disorder; inflammatory bowel disease; multiple sclerosis;
KW asthma; osteoarthritis; psoriasis; rhinitis.
XX
OS Homo sapiens.
XX
PN WO200061750-A2.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000MC-US09277.
XX
PR 09-APR-1999; 9905-0128492.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Phan H, Courtneidge SA;
XX
DR WPI; 2000-679486/66.
XX
DR N-PSDB; AAC66076.
XX
PT Novel tyrosine kinase substrate polypeptides and polynucleotides, used
PT to diagnose and treat diseases such as rheumatoid arthritis,
PT atherosclerosis and cancer.
XX
PS Claim 2; Fig 5; 120pp; English.
XX
CC This invention relates to isolated, enriched or purified nucleic acid
CC molecules represented by sequences AAC66076-066079 which encode tyrosine
CC kinase substrate (Tks) Tks 107/GRUB, Tks 113, Tks 118/Dresh or Tks 202
CC represented by sequences AA185660-185663. The Tks proteins are
CC substrates for cytoplasmic tyrosine kinase Src. The invention encompasses
CC probe sequences, recombinant cells, antibodies specific for the Tks
CC proteins, and hybridomas producing the antibodies. The Tks proteins
CC exhibit antirheumatic; antiarthritic; antiarteriosclerotic;
CC immunosuppressive; cardiact; cerebroprotective; neuroprotective;
CC cytostatic; antiinflammatory; osteopathic; antipsoriatic; antiallergic;
CC and antiasthmatic activities. The tyrosine kinase substrates are used to
CC treat and diagnose diseases or disorders selected from Rheumatoid
CC arthritis, atherosclerosis, autoimmune disorders, organ
CC transplantation, myocardial infarction, cardiovascular diseases, stroke, renal
CC failure, oxidative stress-related neurodegenerative disorders,
CC cardiovascular disorders, cancer, and immune-related disorders selected
CC from chronic inflammatory bowel disease, chronic inflammatory pelvic
CC disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, and
CC rhinitis. The proteins can also be used to identify modulators of its
CC activity. DNA sequences encoding the proteins may be used in gene therapy
CC methods, and to produce transgenic animals.
XX
SQ Sequence 1519 AA;

```

```

Query Match 2.7%; Score 143; DB 21; Length 1519;
Best Local Similarity 19.7%; Pred. No. 0.0038;
Matches 190; Conservative 87; Mismatches 284; Indels 402; Gaps 41;

```

```

QY 12 LTLGPRRADSEFQALLDIFPEKPLPTAFVDTG---EALLLPDMKLKIRISEVL 68
DB 549 LRLTTPP-----CPPEPPSPSDTLNTLHYLSLRD----- 582
QY 69 RVDAAIODELPQOLLFVOSFGIPVSMKSLQFLDQA-----VAHDPOTLE-- 116
DB 583 -----LQTLGLSVLRLQAPPLPALIPALSOLOQSGDPPVQLLLIHDDLPLELC 636
QY 117 -----QINIMDKNYM-----AHLVEVOHE-----RGASGC--Q 141
DB 637 GFGCAEVLSENDLKRVAKPELOMELGHRDPPSSHVEIHQEVVRLCRICQGLSVRQ 696
QY 142 TFHSLTASLPPRRDSTEAPKPSPPQIGQGRIRVGTQLRVYGPEDDLAGMFLQIFPL 201
DB 697 AIELLEGAAPEEPEEAVGMKPIQ-----KVLADPRLTALQRD--GGAILLRLRS 744
QY 202 SPDRMOSSSPRPVALALQALGQELARVVGSPFEVPGITVRVQLATILSSPHGALV 261

```



```
Db 745 TPSSKLEGGPATLYQVEYDEATHQ-LVR-----LSNLH----- 776
QY 262 MSMHSHFLACPLRLQLOXYQCVPODGFSSFLKVLQMLQWLDSPVE-----G 313
Db 777 -----VOQGEQROC-----LRLQOVLQWLSPGEEQLASFAMPG 811
QY 314 GPLRA-----QLRMLASQASAGRRLSDVR----- 337
Db 812 DTLSLQETELRFRAFSAEVQERLQARALALEENATSOQKVLDFEQRLQVEESGLHRA 871
QY 338 -----GGLRLA-----EALAFRODLEVVST--YRAVI----- 364
Db 872 LRLQFFQOAHFWVEGFARLAGAGREAVLALALRRAPESAGTFQEMRALALDLSG 931
QY 365 -ATLRSGEQCSE-PDLISKVYLQGLIEVRSPLLELTAFFSATADAA-----SP 412
Db 932 PAALREMGRCQRCQELERRIOQHLGEASPR-----GYRRRRADGASSGQWGPSP 985
QY 413 FPAKPVVVSLLQEEEPILAGKPGADGSL-----EAVRLGPSS-----GGL 457
Db 986 SPS-----LSLLPSS---PGPRAPSHCSLAFCGEDYEEGPELAPAEGRPPRAVL 1036
QY 458 VDWLEMLDPEVYSSCPDLQLRLFSRRKKGQAVSFRPYLLTLFTHOSSWPTLHOCIR 517
Db 1037 IRGLEVTSTEVVD-----RTCSPREH----- 1057
QY 518 VILGKSREORFDPBSLDFLACIHVPRIWGRDORTPOKRRELVLRVQ----- 567
Db 1058 VILGRAR-----GPDGPMG-VGAPRMRKRISIAOQRLVSELJAGEQDYVATLSEPV 1108
QY 568 ---GPGLISLVELLILAEAFTRSODGDTAACSLLIQARL-----PLLS 606
Db 1109 PPPGPBL-----TPELRGTMAAALSARELRSFHRTHFLRELQCATHPL--- 1153
QY 607 CCCGDDESVRKVTHELSCGIQWGDVILGRRCRDILLQYLQRPBLRVVPPEVLLHSEGA 666
Db 1154 -----RIGACFLRHGDOF-----SLYAQYVKHRHKTENGLALSPLSKGS 1193
QY 667 ASSSVCKLDGL-----IHRFTTLADTSDSRALENRGADASMACKLAVANHPILL---L 717
Db 1194 MEAGPYLPRALOQPLEQLFRIGRLLE-----LLREAGPELSSCCRALGAANVQLREDEA 1248
QY 718 RHLPMIAALHGRTHLNFQEFROQNHLSCLVLGLLELLOPHVFRSEHOGALMDCLLSF 777
Db 1249 RGRDILAVEAVRGEIDLKEQGQLHRDPFTYICGRKKCIR-HVFLFEH-----LLLF 1300
QY 778 IRL 780
Db 1301 SKL 1303
```

Search completed: May 7, 2003, 19:40:57
Job time : 55 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:39:33 ; Search time 27 Seconds
(without alignments)
3663.790 Million cell updates/sec

Title: US-09-929-769-7

Perfect score: 5241

Sequence: 1 MHILVHAMVILLTLGGPPRA.....MDPSAQISEALRIILHMEAVM 1029

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3244	61.9	640	2 T08758	hypothetical prote
2	1240.5	23.7	310	2 T46429	hypothetical prote
3	156	3.0	2472	2 E83594	still frameshift p
4	154	2.9	1638	2 T30313	chemotaxis protein
5	149	2.8	1217	2 T00607	hypothetical prote
6	147.5	2.6	878	1 RRSIB	RNA-directed RNA p
7	134.5	2.5	2535	2 T04824	hypothetical prote
8	131.5	2.5	882	2 AE0119	Clp Arpase (import
9	131.5	2.5	1386	2 T00257	hypothetical prote
10	129	2.5	1626	2 A75613	hypothetical prote
11	128	2.4	1012	2 I53172	RAE-28 - mouse
12	128	2.4	1021	2 H75423	hypothetical prote
13	126	2.4	880	2 AE0179	hypothetical prote
14	125.5	2.4	1196	2 T13057	KIAA0729 protein -
15	125.5	2.4	1211	2 C83110	probable exonuclea
16	125	2.4	949	2 T44577	hypothetical prote
17	123	2.3	939	2 G75481	sensory box/GGDEF
18	122.5	2.3	738	2 A53542	breifeidin A-sensit
19	122.5	2.3	1366	2 T35985	probable large pro
20	121	2.3	1134	1 A35955	meta-vinculin - hu
21	121	2.3	1135	1 A39997	meta-vinculin - ch
22	121	2.3	1239	2 S74355	hypothetical prote
23	120.5	2.3	882	2 T43250	spindle pole body-
24	120.5	2.3	1111	2 T01239	hypothetical prote
25	120	2.3	555	2 H83043	hypothetical prote
26	120	2.3	2698	2 B96671	similar to transla
27	120	2.3	3149	1 OQB8	Bip1 protein - hu
28	119.5	2.3	818	2 S62790	mismatch DNA recog
29	119.5	2.3	2048	2 C84609	hypothetical prote

30	119.5	2.3	2895	2 H85362	hypothetical prote
31	118.5	2.3	920	2 H82302	ATP-dependent heli
32	118.5	2.3	914	2 T17233	hypothetical prote
33	118.5	2.3	1237	2 A31789	band 3-related pro
34	118.5	2.3	1237	2 A56764	band 3-related pro
35	118	2.3	1136	2 H95654	F508.33 [imported]
36	117.5	2.2	2388	2 JEO271	beta spectrin, bet
37	117	2.2	1066	2 T10108	vinculin - mouse
38	117	2.2	2442	2 T08621	centrosome associa
39	116.5	2.2	1234	2 A34911	band 3-related pro
40	116.5	2.2	1538	2 B90924	probable ATP-depen
41	116.5	2.2	1538	2 F85772	ATP-dependent heli
42	116.5	2.2	2541	2 S11661	talin - mouse
43	116.5	2.2	2777	2 D96746	hypothetical prote
44	116	2.2	465	1 WMBE28	UL38 protein - hum
45	116	2.2	5032	1 A35041	ryanodine receptor

ALIGNMENTS

RESULT 1

T08758 hypothetical protein DKFzp586J0619.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999

C:Accession: T08758

R:Wandutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z16471

A:Accession: T08758

A:Molecule type: mRNA

A:Residues: 1-640 <MAN>

A:Cross-references: EMBL:AL050110

A:Experimental source: adult uterus; clone DKFzp586J0619.

C:Genetics:

A:Note: DKFzp586J0619.1

Query Match	Score	DB 2:	Length
Best Local Similarity	100.0%	Pred. No. 4e-222;	
Matches 633; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY 397	ELTTAFSSATADAAAPFPACKPVVVVSLDLEEBPLAGCKPGADGSLAAVHLGPPSGI	456	
DB 8	ELTTAFSSATADAAAPFPACKPVVVVSLDLEEBPLAGCKPGADGSLAAVHLGPPSGI	67	
QY 457	LYDMLMDPEVYVSSCPDQLRLPFRSRKKGQAOVPSRPYLLTFTQSSPPTLHOCI	516	
DB 68	LYDMLMDPEVYVSSCPDQLRLPFRSRKKGQAOVPSRPYLLTFTQSSPPTLHOCI	127	
QY 517	RYLGLRSRQRFDPASLDFLMACIHVPRIWGRDQRTQOKRREELVLRVQGHLSLVE	576	
DB 128	RYLGLRSRQRFDPASLDFLMACIHVPRIWGRDQRTQOKRREELVLRVQGHLSLVE	187	
QY 577	LILAEETRSDQDPAACSLIQARLPLLSCCGDDESVRKYTEHLSSGCIQMGDSVLGR	636	
DB 188	LILAEETRSDQDPAACSLIQARLPLLSCCGDDESVRKYTEHLSSGCIQMGDSVLGR	247	
QY 637	RCRDLLQIYLQRPRLRVVPEVYLHSEGAASSVYKLDGLIHRFTLLADSDSALEN	696	
DB 248	RCRDLLQIYLQRPRLRVVPEVYLHSEGAASSVYKLDGLIHRFTLLADSDSALEN	307	
QY 697	RGADASMACKLAVAPHLRLRLPMTALNGRTILNFOEFRQOHLSCFTLVGLLEL	756	
DB 308	RGADASMACKLAVAPHLRLRLPMTALNGRTILNFOEFRQOHLSCFTLVGLLEL	367	
QY 757	LQPHVFRSEHQALWDCLLSFIRLLINRKSSRHLLAFTNKVFQFIHKYITYNAPAISE	816	
DB 368	LQPHVFRSEHQALWDCLLSFIRLLINRKSSRHLLAFTNKVFQFIHKYITYNAPAISE	427	
QY 817	LQKHADPLHDLSPFNSDLYMLKSLAGLSLPRDQTDGGLDEEGEESASGLPLVSVS	876	
DB 428	LQKHADPLHDLSPFNSDLYMLKSLAGLSLPRDQTDGGLDEEGEESASGLPLVSVS	487	

QY 877 LFTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRRPELISFSTNLQRLMSSAE 936
 DB 488 LFTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRRPELISFSTNLQRLMSSAE 547
 QY 937 CCRRLAFLSLRSMQNSPSTAAAFLEPFMYCLGSQDFEVYOTALRNLPETALLCOEHAAY 996
 DB 548 CCRRLAFLSLRSMQNSPSTAAAFLEPFMYCLGSQDFEVYOTALRNLPETALLCOEHAAY 607
 QY 997 LHRAPFLVGMVGMQMPDPSAOISEALRIHMEAVM 1029
 DB 608 LHRAPFLVGMVGMQMPDPSAOISEALRIHMEAVM 640
 RESULT 2
 T46429
 Hypoetical protein DKFp434C0126.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46429
 R:Ansoyge, W.; Wiltner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, January 2000
 A:Reference number: 223028
 A:Accession: T46429
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-310 <AAA>
 A:Cross-references: EMBL:AL137358
 A:Experimental source: adult testis; clone DKFp434C0126
 C:Genetics:
 A:Note: DKFp434C0126.1
 Query Match 23.7%; Score 1240.5; DB 2; Length 310;
 Best Local Similarity 92.6%; Pred. No. 2.1e-80;
 Matches 249; Conservative 2; Mismatches 7; Indels 11; Gaps 1;
 QY 756 LDPHVRSEHOGALMPCCLISFTRILLNKRSSRHIAAFINKFQFIHKYTYNAPAAIS 815
 DB 1 LDPHVRSEHOGALMPCCLISFTRILLNKRSSRHIAAFINKFQFIHKYTYNAPAAIS 60
 QY 816 FLQKHADPLHDLSPNSDLVMLKSLAGLSLPSRDRTRDRLDEGESESSAGSLPLVSV 875
 DB 61 FLQKHADPLHDLSPNSDLVMLKSLAGLSLPSRDRTRDRLDEGESESSAGSLPLVSV 120
 QY 876 SLEPTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRRPELISFSTNLQRLMSSAE 935
 DB 121 SLEPTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRRPELISFSTNLQRLMSSAE 180
 QY 936 ECCRNLAFLSLRSMQNSPSTAAAFLEPFMYCLGSQDFEVYOTALRNLPETALLCOEHAAY 995
 DB 181 ECCRNLAFLSLRSMQNSPSTAAAFLEPFMYCLGSQDFEVYOTALRNLPETALLCOEHAAY 240
 QY 996 VLLHRAFLVGMVGMQMPDPSAOISEALRIHMEAVM 1029
 DB 241 VLLHRAFLVGMVGMQMPDPSAOISEALRIHMEAVM 640
 RESULT 3
 E83594
 still frameshift probable component of chemotactic signal transduction system PA0413 [1m
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: E83594
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mlczoguchl, S.D.; Warren, P.; Hickey, M.J.; B.
 adman, S.; Yan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lm.
 :; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950; MOID:20437337; PMID:10984043
 A:Accession: E83594
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2472 <STO>

A:Cross-references: GB:AE004478; GB:AE004091; NID:99946261; PIDN:AA03802.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA0413
 Query Match 3.08; Score 156; DB 2; Length 2472;
 Best Local Similarity 19.1%; Pred. No. 0.042;
 Matches 229; Conservative 142; Mismatches 386; Indels 440; Gaps 53;
 QY 8 AMVLTLLGPFRADDSFQALDLDIWPPEKPLPAFLVDNSE--EAL--LLPDV----- 57
 DB 846 AQNVPSVLEPPADE-----FYDELRVFLIEEAGEVLEIGRLIPANKADHD 895
 QY 58 -----LK--LRMRSEVL-----RLVDAALDLEPQQLLFVQSF 90
 DB 896 REALTEVRAFFHTLKGSRVRAVLIGELAMSIENLFNRLRSIASEPVQ----- 947
 QY 91 GIPVSSKLLQFLDQAVANDPQTLQONIMDKNMALHVEVQHGRGSGGOTFSLTAS 150
 DB 948 -----RVVDQVALLPELVEE-----FAANAQRQRDVDLLAATAHALAKE 989
 QY 151 ---LPPRDSTEARPKSPSEPIG-----OGRIYVGTQLRYLGPEDDLAG 193
 DB 990 PLRPPAPDGGVP--PEAGAEQPSLNGVQAPPLADAPQAANAQSDVELDQ----- 1043
 QY 194 MFLQIFPLSPDPRMOS-----SSPRVALAQALQGLARVQGSPEVPGITVR 243
 DB 1044 -LEIFITNEAETHLEALVGLADCARLPQVTDALRAL-----HTLKSAHNAIGIL-- 1095
 QY 244 VLQALATLLSPHGALVMSHRSFLA---CRLRQIQVQRCVPQDTGFSFLKY 298
 DB 1096 PLAIATPPEK-----LVREKSNLAFDLRELLHDEQLRIGLEVOYGAORPLNPI 1149
 QY 299 -----LIQMLQWDSPGV-----EGGPLRAQLBMLASQASGRRLSDVRGLRLA 344
 DB 1150 PGSDALLERIEALHQRHSLAEARYSDAGERRPLIEAFVYGMQILDAEDLLRWH 1209
 QY 345 EALAFRODLEVYSSVAVATLRSQ-----EQCSVEPLISKVLOGL----- 388
 DB 1210 EHPQERDEL---SALREELTLDGRARHMLPOVEELCAQLALVAYVEGRLAVSPAF 1265
 QY 389 --EVRSPH-----LEELTA-----FFSATDAASPPF 414
 DB 1266 FEERQAHLEALIGMDVAAAGLOVTPPRERVAALQELIEBAPAAVPIIDPESIGADDF 1325
 QY 415 -----ACKPVVVVSSLLQEEEPAGKPGADGSL-----BAVRLGPSSG--LLV 458
 DB 1326 PEDEPALPEAVYEAGAPAEETVPAPAPAPGRELDEBWSIFLEAVDILBSAGALA 1385
 QY 459 DWLEMLDPEVYSSCPDQLRLRSRRKGGOAVPSFRPILTLFTQSSWPTLHCICIR 518
 DB 1386 QW--QAEPGALSSLSAQ--RLHLTLKGARMAETAEIGD-----LAHE-----LEA 1428
 QY 519 LLKSGROFPDPSALDFLMACIHVPRTWGRDRT---PQKREELVLRV---QGP-- 569
 DB 1429 LYEGLDVRRIOHSPQLGLGLQACH--DRLAEQLDLSAGQPLADPHDLIOSTRRRQGPVA 1487
 QY 570 -----ELISVELILAEATRSQDGTAACTLQIARLPLLLSCCGDDESVRKVTLS 623
 DB 1488 EAATPGAESEVVELVAPA---VEPAPAPAAEAEBNDPELVEFL--EEGFLIDLSAA 1541
 QY 624 GCIOQMGDSV----- 633
 DB 1542 AALQRMWDVDVNTIELEALORDHTTLKGARMAEIGEIGDLAHELELYEGICGRLRAS 1601
 QY 634 -----LQRRCDLLLOLY-----LQR-----BELAVPYEVU-----LH 662
 DB 1602 PALFGLQORCHDELAELAEVGRHTLPDGOALIAETIRSRSDDELVSPTSVSLKPLA 1661
 QY 663 SEGAASSVCKLDLIRFTILLAD-----TSDSRALENRQADASMACKRLAVA 711
 DB 1662 AKGAADSEILDTFLLEADDLLELEALGRMDGNGDQAPLDD----- 1706


```

Db 588 --DKRRLCSAAYPKGVENKSL-----KSKVIGEQAYKVVYEALRLVGWMPYPLNK 638
OY 607 CCCGDESVRYKTE-----HLSCIOQMGD-SVIGRRCRDLLLOLYLOR---DELRYVPV 657
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 639 ACKNNAARHLLEAKGPFLLDEFLEAMSELSEFGTEFEGFNKIKLYTRENLAFLNKVPV 697

RESULT 7
T04824
hypothetical protein F10M23.350 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T04824
R:Bevan, M.; Lecharny, A.; Chedior, F.; Kravitzky, M.; Kreis, M.; Hohnsels, J.; Mewes, F.
Submitted to the Protein Sequence Database, February 1999
A:Reference number: 215385
A:Accession: T04824
A:Molecule type: DNA
A:Residues: 1-2535 <BEV>
A:Cross-references: EMBL:AL035440
A:Experimental source: cultivar Columbia; BAC clone F10M23
C:Genetics:
A:Map position: 4
A:Insertions: 395/1; 613/3; 777/1; 826/3; 1945/2; 1993/3; 2061/3; 2129/3; 2178/3; 2295/3; 2
A:Note: F10M23.350

Query Match 2.6%; Score 134.5; DB 2; Length 2535;
Best Local Similarity 19.9%; Pred. No. 1.5;
Matches 213; Conservative 128; Mismatches 336; Indels 391; Gaps 53;

OY 21 DDFEQAL---LDIFPEEKPLPTAFVDTSEEA-----LLLPMLKL--RMIRSEVRL 70
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 568 EDAEKEYIGIVSDIWSELCSNP-----IDSVEEAEKCFHIKILDSIKIYVRAVPNELGGS 623
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 71 VDAALDLEPQQLLFVQSGIPVSSMKLQFLDQAVAHDPQTELEONIMKNYAHLYE 130
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 634 FDIFFKPLS-----NSFGMPVELQALLSLLEYSISWTGRS----- 659
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 131 VOHERGASGGQTFHSLTLASLPPRRDSTEARPKSPSPQIGRIQVGTQRLVGLPED- 189
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 660 -QSDNG-----PTRRP--PLMHKHLRVINILLPSPHNG 690
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 190 --DLA-----GMFLQIF-----PLSPDPRMOSSPREVALA 218
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 691 VKDLAVNLAVAAINSTGAFFENNPSEIGAMFLPCFEKIKLLELOEAQVSSVVSFL 750
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 219 LQ--GALGQELAR--VVOGS--PEVPGIVR-----VLQALATLLS-----PHGA 259
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 751 CDAVSTVGNNLFKHMWDIVRSSLSHKGVISGFSPPLITCLLQKCVRLNSESSTSLPEKSA 810
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 260 LVMSNHRSHFLACPLLRQLQYQRCVPODTGFSFLKVLQMLQMLDPPGVEGGLRAQ 319
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 811 ISL-----YVCSFLKLYLQIQ---VDSKLISCLQSVLSEVDESKSLCEWRRLR-- 858
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 320 LRLMASQASAGRLSDVNGGLR-----LADALAFRODLEVASTVRAVIAITLRSGECC 373
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 859 MLICSSQS-----LSNERPIILHSRTTGLPADSSFAETLDEIKRLVR----- 901
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 374 SVEPRLISKVLOGGLEVRSPHLELLTAFFSATADAASFPRACKPVVVVSSLLQEEPL 433
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 902 SISPELIGIYKAFSSALICATPESITLQFASVMQSNAFY--TPPSFQSTITPLEENFL 960
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 434 AGKRGADGSGLEAVRLGPPSSGLVDWLEMLDPEVAVSCPDQLRLLEFRRKGCAQVP 493
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 961 GN-----LSKLSPLDLPASSSE----- 976
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 494 SFRPVLTLFTHQSSWPTLHQCIRVLLGKSRQRFDPSSASLDLFLACIHWPRIMQGRDR 553
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 977 -----FTGS-----GNLCGYVD--SEIDF-----SGHS-- 998
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 554 TPQKRREELVLRVQGPRLISIVELLAEAFPSQDGPDAACSLIQARLLPILSCCGDGE 613
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 999 -----SYTEELRSKMDNDESSAFSIFLKAAPPVLLINAIMSMDI 1039
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

```

```

OY 614 S-----VRKTEHLSCIOQMGDSVIGRRCRDLLLOLYLORPELRYPEVLL 661
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1040 SCLPERRISELLILKVSQPKSGSI-----DS-----NIQILFLPQIRSSYKQ--PAPVL 1090
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 662 HSEGAASSVCKLIDGIIHFTLLADTSDSRALENQAD--ASMCRRLAVAHPLLRL 718
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1091 HO-----LSEIC-----LRLMKNLFSQISEPELVSGSSNKLPAFAFKHQAWEVLVC-- 1139
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 719 HCPMTAALLH-----GTHLNFQEFQONHLSCLFHLVGLL 754
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1140 H--PVVWALLESPDCCGTPPVQVNEIFSETSLTMR--LVFSEIDQ-----HILNLL 1188
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 755 ELQRPVFRSEHOGALM-----DCLLSF-----HLLNLYR-----KSSRL 791
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1189 VSTCEFLFDEKRPNNKMKDKLKNKSIATFKDLYELLLEFRKFLCVCSSQSYSLQD 1248
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 792 AAFINKVQFIHKYTYVNAAPAIISFLQKHADPLHDSFQNSDLVMSKL--LAG-----LS 845
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1249 AOLIHALTRFISPFKLFN--IAHSMCLKIDE--EGLTSPNSSIILSLGLIGAGAFEMLV 1304
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 846 LPSRDRTRDRCGDE-----EGEESASAGSL-----PLVSYLFTPLTAAEMA----- 887
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1305 LVSHQFTARKGVYDILWLELEEKVNASNILEKYYSMAKCFSTSIDDSADICLKVCGGIF 1364
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 888 -----PYMKRLSR--GQTVEDLLEVLSDIDEMSRREPEILSF 922
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1365 RCKHNNVAVNDPLVNLKISLVIGRTPEDL--IHCIRASITRAKILFY 1410
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

```

RESULT 8

AE0119
Clp ATPase [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AE0119
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.
II, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; M01D:21470413; PMID:11586360
A:Accession: AE0119
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-882 <KUN>
A:Cross-references: GB:AL590842; PIDN:CA089816.1; PID:q15979042; GSPDB:GN00175
C:Genetics:
A:Gene: clpB3
C:Superfamily: endopeptidase Clp ATP-binding chain

```

Query Match 2.5%; Score 131.5; DB 2; Length 882;
Best Local Similarity 21.3%; Pred. No. 0.5;
Matches 170; Conservative 95; Mismatches 279; Indels 253; Gaps 36;

OY 22 DSEFQALDIWPEEKPLP-----TAFVDTSEBALY-----PMLKLRMR 64
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 220 DSETRQMDILSRKRNKNLIVGEPGVKTALY-----EGLALRIAGNVPPSLKTVSVR 274
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 65 SEVLRLYDA-----ALQDLEPQQLLFV-----OSFG----- 91
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 275 TLDLGLQAGAGVKGFEQRLKNVLEAVQO--SPSPVLEFIDEAHTLIGAGNOAGADAAN 333
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 92 -----IPVSSSKLQFL--DOAVAHDPQTELEONIMKNYAHLYEVEHG 136
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 334 LKLPALGELKRTIATFTWSEYKQYFERDALERRFQMKVDEPDDDTACMLRGLKERY 393
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 137 AS--GGQTFHSLTLASLP--PRDSTEARPKSPSPQIGQ--RIR-----VGT 180
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 394 ATHHGVHILDAITAAVTLSRRFLTGRQLPDKAAVLDLDTAGARVRSIDTLPALMAINA 453
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 181 QLRVLGPEDDLAGMFLQIFPLSPDPRMOSSPREVALAL--QALGQELARVVQGSPEVPG 239
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

```

Db 454 ELAALAMEGQAIEDLLLPVNSTRLPEIEORRAALAVQHTLEQY-----KEEKR 506
 QY 240 ITVRVQALATLLSPHGALVMSHRSHFLACPLLRQLCOYRCVPQDTGFSSLPKVL 299
 Db 507 LTTLLIARQDIANVAH-----LVALQGEIVQIQGNAP-----LLSLDVDR 548
 QY 300 LQMLQMLDSDGVEGGLRAQRLMLASQASAGRLSDVGGLLRLAEALFR---QD--LE 354
 Db 549 TVAAVINDMVG---PLSLILK-----DEQTLLOLEHNLATRVVQODPALG 592
 QY 355 VASSTAVATATLRS-----GEOCSVEPDLISKVLOGLIEVRSHLELLTAFFSATADA 409
 Db 593 ALQRLAAATGTLSENGPLGAVFLVGPSSIGKT-----ETALALA 633
 QY 410 ASPFPACKPVVVSLLLOEE---EPLAGKPG---ADGSL-EAVRLGPSGLVDWL 461
 Db 634 DSLFGKSLITINLSYEAMHVSQKSPGVGVGGVLTENVARKRYSVLLDEV 693
 QY 462 EMDPEVSSCPDLQALLSR---RKKGQAOVPSFR-----PYLLTFTHQ 506
 Db 694 EKAHRDLN-----LFYQVDFRGFMDEGRE--IDFRNVILMTANIGSDHMLQLLDEQ 746
 QY 507 --SSMPTLHOCIRVLLGKSREORFSPASLDFMACIHPVRIMQGRDQRPQKRREELVL 564
 Db 747 PEATHSDHLHLRPL---RDRQPALARF-----QTLIT 779
 QY 565 R-VQGPGLISLVELLAEATRSQDGDPAACSLIQARLPILLSCCGDDSEVRKVTLS 623
 Db 780 RPLNAIARIVIEKMLQAVAKRLNKHGHLCTIESLYDTLVACLLPDTGARNI----- 834
 QY 624 GCQQQWDSVIGRRCDLLQVYQRELVNPVEVLHSEGAASSVCKLDGIHNFIT 683
 Db 835 -----DSLNGQILFVLSQQLLR-----MSEQQRITS-----LT 864
 QY 684 LLADTSDSRALENAGD 700
 Db 865 LGWDEADGITLEPEGGE 881

RESULT 9
 T00257
 hypothetical protein KIAA0476 - human
 C:Species: Homo sapiens (man)
 C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
 C:Accession: T00257
 R:Seq1, N.; Ohita, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.; Nomura, N.;
 DNA Res. 4, 345-349, 1997
 A:Title: Characterization of cDNA clones in size-fractionated cDNA libraries from human
 A:Reference number: Z14085; MUID:98116662; PMID:9455484
 A:Accession: T00257
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1386 <SEK>
 A:Cross-references: EMBL:AB007945; NID:g3413913; PIDN:BA32321.1; PID:g3413914
 A:Experimental source: brain
 C:Genetics:
 A:Note: KIAA0476

Query Match 2.5%; Score 131.5; DB 2; Length 1386;
 Best Local Similarity 22.0%; Pred. No. 0.98;
 Matches 183; Conservative 71; Mismatches 315; Indels 263; Gaps 37;

QY 76 QDEPQOLLFVQ--SFCIVSSMSKLLQFLDQAV--AHDPQTEONIMDKNVAHLEV 131
 Db 510 QLHTQMFOSFIECSFG--SARHALEFFDSCVEKVHPEQ-----EKPEPTPLVEL 559
 QY 132 QHERGASGGTFFSLTASLPPRDSREAPKPKSSPEQPIGGGIRVGTOLRVLGPPDDL 191
 Db 560 EELISG-----SELVFTTP--PEEPALPESSESTP-----QTCYDGPPELR 598
 QY 192 AGNFLQI-----FISPPDRMOSSPPRYALALQALGELARVVGSEVPICITRVL 245
 Db 599 AELFESIQEQPGALPVGPGRSAPSSAPARTKQEMKVAQ---RMAQKSAAVPELMARCL 655

QY 246 QA-----LATLLSPHGALVMSHRSHFLACPLLRQLCOYRCVPQDTGFSSLP 295
 Db 656 LGHCYGLMFLCPLPAYVRSAP---SRVQALHTAYH---VLRQMESGVVLPEDEVY---- 704
 QY 296 LKVLQMLQMLDSDGVEGGLRAQRLMLASQASAG-----RLSLDR 337
 Db 705 -RVLMQCSHYGPVL-----SVRVLEMQRQAGIVPNTTYGYKRAVLESKWPSGTP 756
 QY 338 GGLRLAE-----ALAFRODL-----EVVSVRAVIATLRSGEQCSVEPDLISKVLOG 386
 Db 757 GGRIRMAKLRNVVLGAQFROPPLREQQQQQQQQQQQQQQQQQQQEQVSAHQEAGSSQAP 816
 QY 387 LIEVRSP--HLEELLTAFFSATADASP----- 412
 Db 817 YLERPSPLRLQRTQTWAGRLRDPASPPGRLVKSLSGARQAQPTVEAGVAHMEALG 876
 QY 413 --FPACKPV-----VSSLLQEEEPPLAGKPGADGGSLEAVRL-----GSSGL 457
 Db 877 VLEPRGSPVPMHDSLSDLTGEPLPGSGPGGSALSQSTEALEGLSGRQPKAGR 936
 QY 458 VDMLEMDPEVSSCP-----DLQLRLFSRRKKGQAOVPSRPYLLTLFTQSSMPT 511
 Db 937 QD-----EAGTPRGLGARLQQLTPSRHSPASRIPEPLPDLPPARRSPMDSL 987
 QY 512 LHOCIRVLLGKSREQ---RFPDSASLDFMACIHPVRIMQGRDQRPQKRREELVRLVQ 567
 Db 988 LH-----PREPGSTASASISGEWDLSE----- 1013
 QY 568 GPGLISVELLAEATRSQDGDPAACSLIQARLPILLSCC--CGDSEVRKVTENHSGC 625
 Db 1014 ---SSLNLSLRSSERLSD---TPGSFOSPSLEILLSSCSLGRACDSLYDEIYAG- 1065
 QY 626 IQQWG--DSVIGRRCD-----LLQLYQLPELRVPEVVLHSEGAASSVCKLD 675
 Db 1066 ---WAPDSNLTNTCPACACFPVLLSVQTLDSRPS--VSPK-----SAGASGSKDAPV 1116
 QY 676 G-----LIHREFTLLAOTSRALENAGADASMACRLAVALPILLRHPMLAALLHGR 730
 Db 1117 GGPGVLSDRRLCALDEPOLCNGMGASRRVSGAMVLSPLVLRKELESV----- 1170
 QY 731 THLNQFQROONHLSCTFLHYGLLELLQPHYFREHOGALWDCLSFTRLLL 782
 Db 1171 -----ENEGS---EVLALPEL-----PSAHFTFNWNLMTYQRL 1203

RESULT 10
 A75613
 hypothetical protein DRA0166 - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: A75613
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: A75613
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11626 <MHI>
 A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12309.1; PID:g646
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0166
 A:Map position: 2
 C:Superfamily: Deinococcus radiodurans hypothetical protein DRA0166

Query Match 2.5%; Score 129; DB 2; Length 1626;
 Best Local Similarity 20.4%; Pred. No. 1.9;
 Matches 170; Conservative 91; Mismatches 270; Indels 302; Gaps 38;

A:Map position: 1

Query Match 2.4%; Score 128; DB 2; Length 1021;

Best Local Similarity 21.7%; Pred. No. 1.1;

Matches 195; Conservative 74; Mismatches 314; Indels 314; Gaps 40;

QY 121 DKRYMAHLV-EVCHENGASGGQTFHSLTLTASLPDRDSTAPKPKSPEOP---ICGR 175
 DB 70 DKRLLELLRELAQEAFL-----RLDLCAPEQEPAPLPPELELLGRSK 112
 QY 176 IRYGTOLRVLPEDDLQAGMFLQIFPLSPDRMQSSSPRY-----ALAL 219
 DB 113 KRYVQKQF-----LAGV-----RAGRTPLVYVTRGAGCKTVLLDHLAAL 152
 QY 220 QQA-----LQELARVVOGSP-----EVPQITVRLQALATLSS---PHGALVSM 264
 DB 153 QAGARVERLNLGDLALAGPPLRDRSPASLAARQDLALRYLPERLPVGSALLVRY 212
 QY 265 HRS-HF-----LACPLMLQOCYORCVQDQTFSSLFKLVLQMLQW- 305
 DB 213 TENLHFGNDPRLPDGTPATMAAAAEVLK-----RHIF--AGVSYLF--ALEDAAGWP 262
 QY 306 -----LDSPEVGGPRLRAQLMLASQASAGRRL- 333
 DB 263 PGAGELIELOPTLTERATILMARLVSRVADSLAETRNLDRLTLGAGGNARLL 322
 QY 334 --SDVRGGLRLALEALAFRODLEVSSTVAVIATLRSQCSVEPDLISKVLGLIEVR 391
 DB 323 ADHVDV---RLACAAAV---LAVPPTGTRDTASSVPLGDPVPLPAVALALALGGVYAL 375
 QY 392 SPHLEELTAFFSATDAAFPACKRVVYVSSLLQEEPLAGKRGAGAGSLEAVRLG 451
 DB 376 PLIARSLDLE-----VTPPLEEGGPEAGQAVGEOTG--EKYVGM 413
 QY 452 PSSGLVDMLE-LDEEV--VSSCPDLQLRLFSRRKKGQAOVSEFPYLLFTFHQSSV 509
 DB 414 RRPRLP--WIALPLPEAEVATEQVQATQSRDLS-----PEVVPYLLALALAGEM 465
 QY 510 PYTHOCIRVLGKSGREOPDPSASLDFLWACIH-----VPRIMOGDRPT 554
 DB 466 SMLEHVRRA-----RPDDARYLLPALMRIRIAGASSPEREGARAVVTHVASRGEYHA 517
 QY 555 POKRRELYRVQGPPELISVELIAEETRSDODGTAASLIQARPLLLSCCGDES 614
 DB 518 PAARDALFVLEBGAHAHARVKLAESSLDAGNFFETAGTOLEKAL- 564
 QY 615 VRKTEHLSGCIQOWGDS-VLGRRCRDLILQLYLQK-----PEL 652
 DB 565 -----EHLVSQPGSGWGDWRLAOSGLVQALARWRGMAGATEAVSPRTAQSFPRA 619
 QY 653 R-----VVPYEVLLHSEGAASSVCKLIDGLIHRTLLADTSD 690
 DB 620 RLMRGIVADKAGHWEALDHLGAVPGSSPLSTRARYQEGDRLR--RLGCPVAYALALSD 677
 QY 691 -SHALENRGADAMACRKLAIVAHPLLLRLPRITALLHGRTHLNPCEFFQOHLSEFLH 749
 DB 678 AARRLAAGANPEQARVILARA-----ATARRRGYAE---DALRLFGQ 718
 QY 750 VLGLLELLOPHVRSEHOGALMDCLLS-FTRLLLNKRKSSRHLA-----792
 DB 719 ALALL-----PTDLRSADAVLQARLISEQVPLVLLALGRPDALAAQAAQALTLRGEGR 774
 QY 793 APTNKRKYVQFH-----KYTTYNAAPAAISFLQKADPLHDLSFNSDLSVMLKSLAGL 844
 DB 775 AEVTYEVRRHYRVALAVLYLRG--RGVPYLDQPGGP-----EFDTPDLVHARALIAEL 825

RESULT 13

AE0179

Probable ATPase chain of ATP-dependent proteinase YPO1471 [imported] - Yersinia pestis
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
 C:Accession: AE0179

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barral
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MIMD:21470413; PMID:11586360

A:Accession: AE0179

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-880 <KUR>

A:Cross-references: GB:AL590842; PIDN:CMC90296.1; PID:q15979515; GSPDB:GN00175

C:Gene: YPO1471

C:Superfamily: endopeptidase C1p ATP-binding chain

Query Match 2.4%; Score 126; DB 2; Length 880;

Best Local Similarity 21.6%; Pred. No. 1.2; Mismatches 271; Indels 230; Gaps 34;

QY 22 DSEFQALLDIWFPEKPLP-----TAFIVTSEALLF-----PDWIKLRMR 64
 DB 220 DSEIRQIDILSRKNNPILVGEVGYKTALY-----EGALRIADGNVPDSLKTVSVR 274
 QY 65 SEVRLVDA-----ALQDLPEQQLLFV-----QSTG----- 91
 DB 275 TLDGLLQAGAGVKEFEQRLKNVIEAVQO--SPSPVLLFIDEAHTTIGAGNAGGADAAN 333
 QY 92 -----IPVSSMKLQFL--DOVAHDPTLEONIMDKVMHMLVEVCHERG 136
 DB 334 LKRLALRGELRTIAATWSEYKQYFERDAALERRPQWVYDEDDDTACMLRGLERY 393
 QY 137 AS--GGQTFHSLTASLP--PRDSTEAPKPKSPEOPIGG--RIR-----VGT 180
 DB 394 ATHGVHILDAATIAAATLAAVTLNRFLTQRLQDPKAVDLDTGAFARMSIDTLPTALMEINA 453
 QY 181 QLRVLGPEDDLQAGMFLQIFPLSPDRMQSSSPRYVAL-QLALGQELARVVOGSPVPG 239
 DB 454 ELAALAMEQQAIEDLILLPVSSRLPEIQRAALAVGCHTLEQY-----KEEK 506
 QY 240 ITVRVLOALATLLSPHGALVMSMRSHFLACPLRLQOCYORCVQDQTFSSLFKVL 299
 DB 507 LTTPLIEARQDIANAAN-----LVALQELVIOGNAP-----LSLDVIDR 548
 QY 300 LQMLQMDSPGVEGCPRLAOLRMLASQASAGRRLSDVRGGLRLALEALFR--QD--LE 354
 DB 549 TVAAVIVDMWGV--PL-----GSLKDEQTMLSLENRLGERVIOQDAALG 592
 QY 355 VVSTVAVIATLRS-----GEOCSVEPDLISKVLGLIEVRSPHLELLTAFPSATADA 409
 DB 593 ALAQRLEAAKTGLTSENGPLGVLVLPVSGTGT-----ETALALA 633
 QY 410 ASPFPACKPVVNVSSLLQEE--EPLAGKPG-----ADGSL-EAVRLGSSGLVDMWL 461
 DB 634 DSEFGKSKSLITINLSYEQEAHIVSOLKSPGVGVGGGVLEAVRKRYSVLLDEV 693
 QY 462 EMLDPEVSSCPDLQRLFSR--RKKGQAOVPSR-----PYLLTFTHQ 506
 DB 694 EKAHRDVLN-----LFQYQVDRGFMDEGRE--IDFRNVIVTANLIGSHLMQLDDEQ 746
 QY 507 SSWP--TLHOCIRVLGKSGREOPDPSASLDFLWACIHVRIMWGRQRTPOKREELVYL 564
 DB 747 PDADOSTLHELRLPIL--RDHFQPALARF-----QTLTY 779
 QY 565 R-VQGPPELISVELIAEATRSDODGTPAACSLIQAARPLLLSCCGDDESVRKVTHEL 623
 DB 780 RPLQYDALKRITVAIKLNQVAGRLCHGSCQIEDSLDTLVAAACLLPDTGARNI----- 834
 QY 624 GCTQMGDSVYGRRCRDLILQLYLQK 649
 DB 835 -----DSLNOQIIPVLSQQLISR 853

RESULT 14

713057
 KIAA0729 protein - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
 C:Accession: T13057; T08741
 R:Nagase, T.; Ishikawa, K.; Suyama, M.; Kikuno, R.; Miyajima, N.; Tanaka, A.; Kotani, H.
 DNA Res. 5, 277-286, 1998
 A:Title: Prediction of the coding sequences of unidentified human genes. XI. The complet
 A:Reference number: Z17595; MUID:99087487; PMID:9872452
 A:Accession: T13057
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1196 <NAG>
 A:Cross-references: EMBL:AB018272; NID:93882178; PIDN:BA34449.1; PID:93882179
 A:Experimental source: brain
 R:Mambuti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence database, March 1999
 A:Reference number: Z16471
 A:Accession: T08741
 A:Molecule type: mRNA
 A:Residues: 279-1196 <NAG>
 A:Cross-references: EMBL:AL050092
 A:Experimental source: adult uterus; clone DKFZp586G0518
 C:Genetics:
 A:Note: KIAA0729; DKFZp586G0518.1

 Query Match 2.4%; Score 125.5; DB 2; Length 1196;
 Best Local Similarity 20.0%; Pred. No. 2.1; Indels 241; Gaps 39;
 Matches 155; Conservative 101; Mismatches 277;

 QY 213 RPVALALQALGQELARVVGSPVEP-----GITVRVLAQALATLISPHGALVMSMR 266
 DB 35 RPNVHNLYLQPGME-----DGSDDMDTSVEDIGRSCTVRFRTLLILNHEGVKPRSKIL 89
 QY 267 SHFLACPLRLQCYQRCVPQDTGSSFLKVLQMLQMLDPSGVGEGRLQRLMLASQ 326
 DB 90 TEYFAF-----LYEPKAKMEESQF-LLSLQALISTWVHFY--MGTK-GPENQVEVLSEE 140
 QY 327 ASAGRLSDVRGRLRLAEALAFRODLEVSVTRAVITLRSGECSVE----- 376
 DB 141 EGEEEEE-----DLSLAEKTRPALEKMLVALLVQSSERLITLSQDMAALTG 197
 QY 377 ---PDLISKVLQGLIEVR-----SPHEELL-TAFESATADASPAPAC 416
 DB 198 KGPEPLFQHIRDG-INIROTCNLIFSLCKYNNRLAEHIVSMFTSTAKLTPEANPF--- 253
 QY 417 KPVVVVSSLLQEEERPLAGKRGADGSS-----LEAVRLPSSGLLVDMLEMDPE- 467
 DB 254 ---FKLLTLMLE-----FAGGPPGMPFPFASYILQRIWEVLEYNPSOCL--DWLAVCTPRNK 304
 QY 468 ---VSSCPDLQRLLEFSRRKGGQAVSPFRPYLLTLFTHOSSW-----P 510
 DB 305 LAHSVNLQNMENWVERFLAH---NYPRVRTSAVILVSLIPNSFRQMFSTRSLHIP 360
 QY 511 T-----LHQCIRVILG-KSRQGRFDSA-----SLDFLMAC----- 540
 DB 361 TRDPLSPDTVVLHQVNVLLGLLSRAKLYDAVHGTTKLVAFESFTYCLISTEKL 420
 QY 541 -----IHPRIINGRQRTPKRR-----BELVLRVQGPBELISLVE 576
 DB 421 MESTYFMQMLNLEPKLEPAITATNNKQALLSFWMYWCADCPENIRLIVQNPVYTKNTA 480
 QY 577 LILAEFRSDQDGTAAACSLIOARLPLLLSCCGGDESRYKTEHLSGCTIOQMGSDVLR 636
 DB 461 FNYLILADDDQVYLFNGKMLPAYVIGILRLCEQSPATFRQLASHON---IOWAFKNLTF 537
 QY 637 RCRLD-----LLQLYL-QRPELRYVPPEVLLHSEGAASSVYCK-LDG-----LTH 679
 DB 538 HASQYPAVEELFNLMQJLIAQRPDMREBELDIKOFKTTTISCYLRCIDGSSCWTTLS 597
 QY 660 RRTTLAOTSRALE--NRGADASMACCKLAVANPRLLLRLPMTAALLHGRTHINPE 737
 DB 598 AF-RILLESDEDRLLVFNRG-----LILMTE-----SFT 627

QY 738 FROQNHLSCELVHVLG-LLELL-----QPHVRESEHOGAL--WDCLLSFIRELLNY 784
 DB 628 LHMHTAEATACHVTGDLVELLSFLSVLSTKSTRYLQKDYKQALIONQERIEFAHLLTL 667
 QY 785 RNS-----SRHLAAEFINKVQFT-HKITYT--NAPAI 814
 DB 688 LMSYSPPELRNACIDVLEKELVLSPP--DELHTLVPELQHNCTYHNSNIPMSL 739

 RESULT 15
 C83110
 Probable exonuclease PA4282 (imported) - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 14-Sep-2001
 C:Accession: C83110
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Lapid, K.; I
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: AB2950; MUID:20437337; PMID:10984043
 A:Accession: C83110
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1211 <STO>
 A:Cross-references: GB:AE004844; GB:AE004091; NID:99950500; PIDN:AA07670.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA4282
 C:Superfamily: sbcc protein

 Query Match 2.4%; Score 125.5; DB 2; Length 1211;
 Best Local Similarity 21.0%; Pred. No. 2.1; Indels 426; Gaps 49;
 Matches 232; Conservative 118; Mismatches 426;

 QY 24 EQCALIDI-----WPEERPLPAFLVDTSEALLLPDMLEKIRSEVLRVDALODL 78
 DB 251 EQALQRLRGQOQWTFEQR-----LLQSEHA-----QGLAEARQADALATER 236
 QY 79 EPQOLLFVQSGTIPYSSSKLLQFLDQAVANDPQLEQNTIMDKNMAHLEVQHERGAS 138
 DB 297 ETLQWLERL---APVGLIERLQLELRLHSEQOQRO-----RTQQA 338
 QY 139 GGQTFSLTLASLPFRDSTEARPKSPPEQPIGGGRIRYQTL-----RVLPEED 190
 DB 339 GTERLOGIARLOEAREQADNHLRQADP-----LEAPQLESEARLETTLEBOE 393
 QY 191 LAGMFLQIFLSPEDPRWQSSSPRPVALALQALGQELARVVGSPVEPGITVRVLAQALAT 250
 DB 394 L-----HRQSNQRAQSDAARQDMDQGRHVARQALQA-ALRDSQALAA 438
 QY 251 LLSS--PRGALVMSHNRSHFLACPLRLQCYQRCVPQDTGSSFLKVL-----LQML 303
 DB 439 LQDAWVTHQGLATFVQRR-----QRALESQALPBEKSLAHAGEPLERL 484
 QY 304 ---QWLDSPVEGGRPLAOLRMLASQASAGR-----LSDVAGGLRLAEALAFRQ 351
 DB 485 QAWMTALHSGSEPDLARLVELRLQDTSLEQOALKEKMQCYLDQAGLARIGEL----- 540
 QY 352 DLEVSVTVRAVATLRSGEQCSVEPDLISKVLQGLIEV-----RSPHEELLTAFFS 404
 DB 541 DQRMVQGE-QALLDLKRGSSQCAEYKAAQALQVRELLQRLRARSASVEQLRGLVD 599
 QY 405 ATADAASPPACKPVVVVSSLLQEEERPLAGKRGADGSLAVRLGPSGLLVDMLEML 464
 DB 600 GEA-----CPVCG-----SQEHP-----YHSEQLLAALGEHD 627
 QY 465 DPEVSSCPDLO-----LRLIFS---RRKGGQAVSPFRPYLLTLFTHOSSWPTLS 513
 DB 628 DQEQVAEBSLELRQTLVIGIRGYSYSORERLNOSQBOQELTQGLAALDRQDQD--TLP 666
 QY 514 OCIRVLLGKSRQGRFDPASLDFLMACIHVPRIWGRDQRTF-QKRREELVLRVQGPBELI 572

```
Db 687 EELRL-----OPSADLEWL-----AQRDLLAGQRQCCQRPDRLLARQRTQOL 732
QY 573 SLVELLIAEAEETRSDQDGTACSLIQARLPILLSCCGDDSVRKV--TEHLSGCIOW- 629
Db 733 ---QOEIRAEETIIQQROQALTEORQ- RYEHQQQVEEDSQOLRPLLSDEHW---QRMQ 784
QY 630 -----GDSVLRGRCDLLIQLYLQR- PELRVVPE-----VLHSEG 665
Db 785 ADPLRTFQALGESIEGRROQARLQIEQRLEKQRCDESSWQKQSDQCRNEARQAE 844
QY 666 AASSVCKLGLIHRFTLLADTSD- SRALENRGADASMCRKLAVAHPILLRLHPMA 724
Db 845 RAQELAEELNGRGAHLGQHRCAQDWQLSLEHAAQAQSAVETLQA--PLDSLREEDL-- 900
QY 725 ALLHGRTHLNFOEFROQNHU-----SCFLHVIGL-----L 754
Db 901 RLAEALHLLQOQRQROQDEFQRLQADWQAMREODNLDSDRLDALLGISEQATQWREQL 960
QY 755 ELLQPHVFR-----SEHQALWDCLLSFIR-----LILNYRKSSRHIAFINKFVQ 800
Db 961 QRLQEEITRQOTLEAEHQQAQ---LLOHRQRPETDREALEDNLRQQRERLASEQAYLE 1016
QY 801 FHKYITYNAPATSFQKHADPLHDLSPQNSDLVMLKSLLAGSLPSRODRTDRIQDEE 860
Db 1017 -----TYS-----QLOADNQRQEQSALAELEERARAEFRNRGRLINE- 1053
QY 861 GEEESSAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTVEDLEVLSDID--EMSR-- 915
Db 1054 -----LIGSSGDKFRRIAQGYNL-DLIVQHSNVQRLQRLARRYL 1092
QY 916 -----RPELISFSTNLQRLMSSAECCRNLAFLSLRSMQNSP-STIA 958
Db 1093 QRGSESELGLVVDTEMGDELRSVYS-----LSGGETFLISIALALGLASMASSKLRIES 1146
QY 959 AFLPTFMVYCGSDQDEEVVOTALRNL 983
Db 1147 LFIDEGFGSLDPESLQLAMDALDNL 1171
```

Search completed: May 7, 2003, 19:42:51
Job time : 42 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:38:42 ; Search time 15 Seconds

(without alignments) 2845.276 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 5241

Sequence: 1 MHILVHAMVLLTLGPPRA.....MDPSAQISEALRIILMEAVM 1029

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	148	2.8	1302	1 ACSC_ACCEXY	P37718 acetobacter
2	147.5	2.8	878	1 RPO1BDVA	P12918 avian infec
3	138	2.6	2248	1 Y539_HUMAN	O60287 homo sapien
4	138	2.6	3674	1 SPCR_HUMAN	O90268 homo sapien
5	133	2.5	1083	1 T2D3_HUMAN	O90513 homo sapien
6	127	2.4	3321	1 PCN2_HUMAN	O91476 homo sapien
7	125.5	2.4	1211	1 SBCC_PSEAE	O91476 homo sapien
8	123.5	2.4	1455	1 FACA_HUMAN	O15360 homo sapien
9	123.5	2.4	2564	1 SPCO_HUMAN	O91254 homo sapien
10	123	2.3	817	1 MURS_THECA	O91476 homo sapien
11	122.5	2.3	738	1 COG2_HUMAN	O14746 homo sapien
12	122.5	2.3	2035	1 EYPL_MOUSE	O90852 mus musculu
13	122	2.3	1013	1 PTPX_MACNE	O02895 macaca neme
14	121.5	2.3	1058	1 GEM4_HUMAN	P57678 homo sapien
15	121	2.3	1065	1 VINC_CHICK	P12003 gallus gall
16	121	2.3	1133	1 VINC_HUMAN	P18206 homo sapien
17	120	2.3	3066	1 ATM_MOUSE	O62388 mus musculu
18	120	2.3	3149	1 TEGU_EBVV	P03186 epstein-bar
19	119.5	2.3	818	1 MURS_THETH	O56339 thermus the
20	119.5	2.3	1115	1 TRC2_CHLRE	O8VXP3 chlamydomon
21	119.5	2.3	2541	1 TAL1_MOUSE	O91476 homo sapien
22	118.5	2.3	1237	1 B3A2_MOUSE	P13608 mus musculu
23	118.5	2.3	1237	1 B3A2_RABIT	P48746 oryctolagus
24	117.5	2.2	1065	1 SPCP_RAT	O94908 ratius norv
25	117	2.2	1065	1 VINC_MOUSE	O64727 mus musculu
26	116.5	2.2	804	1 RSG4_HUMAN	O95924 homo sapien
27	116.5	2.2	1234	1 B3A2_RAT	P23347 ratius norv
28	116.5	2.2	2541	1 TAL1_MOUSE	P26039 mus musculu
29	116	2.2	465	1 VP19_HSV1	P32888 herpes simp
30	116	2.2	5038	1 RYR1_HUMAN	P21817 homo sapien
31	115.5	2.2	2390	1 SPCP_HUMAN	O15020 homo sapien
32	115.5	2.2	3746	1 ACVS_PENCH	P19787 penicillium
33	115.5	2.2	3791	1 ACVT_PENCH	P26046 penicillium

ALIGNMENTS

34	115	2.2	622	1	FACG_HUMAN	O15287 homo sapien
35	115	2.2	914	1	TORS_ECOLI	P39453 escherichia
36	115	2.2	1241	1	B3A2_HUMAN	P04920 homo sapien
37	115	2.2	5065	1	EPPL_HUMAN	P58107 homo sapien
38	114.5	2.2	854	1	CLPB_THETH	O91663 thermus the
39	114	2.2	979	1	PTPN_HUMAN	O16849 homo sapien
40	114	2.2	980	1	PEX6_HUMAN	O13608 homo sapien
41	114	2.2	1208	1	RCO4_HUMAN	O94761 homo sapien
42	114	2.2	2672	1	GCN1_YEAST	P33892 saccharomyc
43	113.5	2.2	1338	1	LHR_ECOLI	P30015 escherichia
44	113	2.2	1286	1	IRBP_BOVIN	P12661 bos taurus
45	113	2.2	3164	1	TEGU_HSV11	P10220 herpes simp

AC	ACSC_ACCEXY	STANDARD:	PRT: 1302 AA.
AC	P37718:		
DT	01-OCT-1994 (Rel. 30, Created)		
DT	01-OCT-1994 (Rel. 30, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Cellulose synthase operon protein C precursor.		
GN	ACSC.		
OS	Acetobacter xylinus.		
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;		
OX	Glucanacetobacter.		
NCBI_TaxID=28448;			
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 53582;		
RM	Pubmed-8083166;		
RA	Saxena I.M., Kudlicka K., Okuda K., Brown R.M. Jr.:		
RT	"Characterization of genes in the cellulose-synthesizing operon (acs		
RT	operon) of Acetobacter xylinum: implications for cellulose		
RT	crystallization."		
RL	J. Bacteriol. 176:5735-5752(1994).		
RT	[2]		
RP	TOPOLGY.		
RX	Pubmed-11544230;		
RA	Kimura S., Chen H.P., Saxena I.M., Brown R.M. Jr., Itoh T.:		
RT	"Localization of c-di-GMP-binding protein with the linear terminal		
RT	complexes of Acetobacter xylinum."		
RL	J. Bacteriol. 183:5668-5674(2001).		
CC	-1- FUNCTION: Required for maximal bacterial cellulose synthesis. It		
CC	may be involved in the formation of a membrane complex for		
CC	extrusion of the cellulose product.		
CC	-1- PATHWAY: Bacterial cellulose biosynthesis.		
CC	-1- SUBCELLULAR LOCATION: Outer membrane (Potential).		
CC	-1- SIMILARITY: BELONGS TO THE ACSC/BSC FAMILY.		
CC	-1- SIMILARITY: CONTAINS 7 TPR REPEATS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: X54676: CAA38489.1:		
DR	InterPro: IPR003921; CellSynth_C.		
DR	InterPro: IPR001440; TPR.		
DR	Pfam: PF00515; TPR; 4.		
DR	PRINTS: PRO1441; CELLSTHASEC.		
KW	Cellulose biosynthesis; Repeat; TPR repeat; Membrane; Outer membrane;		
KW	Signal.		
FT	SIGNAL	1	32
FT	CHAIN	33	1302
FT	REPEAT	64	97
FT	REPEAT	147	180
FT	POTENTIAL.		
FT	CELLULOSE SYNTHASE OPERON PROTEIN C.		
FT	TPR 1.		
FT	TPR 2.		

FT REPEAT 306 339 TPR 3.
 FT REPEAT 340 373 TPR 4.
 FT REPEAT 572 605 TPR 5.
 FT REPEAT 716 749 TPR 6.
 FT REPEAT 750 783 TPR 7.
 SQ SEQUENCE 1302 AA: 138750 MM: CC1E32F0E1F9D794 CRC64:

Query Match 2.8%; Score 148; DB 1; Length 1302;
 Best Local Similarity 20.8%; Pred. No. 0.043;
 Matches 194; Conservative 120; Mismatches 351; Indels 268; Gaps 43;

QY 102 QFLDQAVAHDPQTLQONIMKNMAHL-----VEVOHE-RGASGQTFISLTLASLPPR 154
 DB 85 QSLQSAARLAPDVLLEAGEYQSHGNDAADTQRRHQAAAPGSTYESQLMDLL--- 141
 QY 155 RDSLEAPKPKSPQPIGGRIRGQTLRVLPEDDLAGMFLQIFPLSPDRMOSPPR 214
 DB 142 -----HQALISQDHLAHARSLASGHSDOAVEYQHLF-----NGSTFPTP 181
 QY 215 -VALALQAL-----GGE-----LARVQGSPEYPIGTFVAVLQALATLSSPHIGALVMS 263
 DB 182 SLAVEYQTLGAVSGQAGTAGDGLIRLVKANPS---DFAQLALAQVLTLY-QPGTRMEG 236
 QY 264 MHASHFLACLLRQLQOYOCVPODGFSSLLFLKVLQMLQMD-----SPGVEGG 314
 DB 237 LQR-----LQALQYQSSAPFEAATAE--KSYRQTLISMLFVTPETPLMOKMDAH 285
 QY 315 PLRAQLMLASQASAG-----RRLSDVRGGLRLAEALAFRODLEVSSTVRAVI 364
 DB 286 PSDSALRTHNAEPAGGPPDGALARODGFALNAGRLSAQA-AFQSALMLNK----- 338
 QY 365 ATLRSGQGCVEPDLISKVLQGLEIVNSPHLE-----LTAFTSATADAASPPACKPVV 420
 DB 339 -----DSGALQGL--GLVAMRAGHNEARHYLEDAIADPPKNAHMRPA----- 380
 QY 421 VVSSLILQEEEPPLAGCKRPGDGSLE-----AVRLGSSGLLVMDLE 462
 DB 381 -----LAGMAVGEISGVNRRLIASGTOEAQRMLTLARPQSGEGLV----- 424
 QY 463 MLDPEVVSQCDLQRLLEFSRRKKG-GQAOVPSFRPLTLFTTHOSSMPTLHQCIRVLG 521
 DB 425 ML-----ADLQ-----RSTGQTEAE-----RNRRLTLARNGDPRILMGLARVLMG 466
 QY 522 KSRQRPDPSASIDFLMACIHVPRIWQGRQRTPOKRREELVAVQGPBELISVELLAE 581
 DB 467 EGSES--EANAALLSRLOG-----RYSQVOQOIE-----VSGIMAE 499
 QY 582 AFRSODQDTRAACSLQA--RLPPLLSCCGDESVKRVTEHLSGCIOQMGDSVLGRRCR 639
 DB 500 A-ARTSDSAOKVSLRQAMTKAP-----DDPMLR--INLANALQOQDSSEAN-- 545
 QY 640 DLLLQVLYQRPRLRVPEVLLHSEGAASSVCKLDGLIHRFTILLADTSDSRAL---E 695
 DB 546 --VMRPLLTSPRTPADYQOAILVYASGNGN-----DTLARLLAGLSPPDYSPAIRTIAD 597
 QY 696 NRGADASMACRKLAVAHPIILLRLHPLIALHLNGRTHLNQEPROQNHLSGFLHVLGLE 755
 DB 598 EMALIKADLASRLSNVSPPTLVREALAAPPTGARGVAVADLFRQD-----D 645
 QY 756 LLDPHVFRSEHOGALMCLLSTIRLLLNYSKSRHLAAFINKEVQFHKYITYNAPPAIS 815
 DB 646 MLNAHM-----ALRI-----ASTRNIDLTTEQRLAVYTEVTKKISNPYAAA 685
 QY 816 FLCKHADPLHLSFDNSDLV-----LKSLLAGLSLPSRDRTDGLDEGESESSAG 868
 DB 686 RL---LAPLDGSGTAGSAMSPDQROTLMQLRMGISVAGSDLLNNGG--DQAAAYDHLAP 741
 QY 869 SLPLVVSILTPPLTAAMAYMKRLSGQVEDLLEVLSDIDEMSKRRPELISFTSNLQ 928
 DB 742 ALQADPEATSPKALALAL--YNGRGKYGHALDIDLAVL-----RHNPDLLARQAAVQ 792
 QY 929 RLMSAEECCRLAFSLARSMONSPSIAAFL 961

DB 793 AAANDGKD---NLAMQALQDGVQOOSPMDARSWL 822
 RESULT 2
 RRPO_IBDVA
 ID RRPO_IBDVA STANDARD; PRT; 878 AA.
 AC P12918;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative RNA-directed RNA polymerase (EC 2.7.7.48) (VPI protein)
 DE (RDRP).
 GN VPI.
 OS Avian infectious bursal disease virus (strain Australian 002-73)
 OS (IBDV).
 OC Viruses; dsRNA viruses; Birnaviridae; Avibirnavirus.
 OX NCBI_TaxID=10997;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88160058; PubMed=2831661;
 RA Morgan M.M., Macreadie I.G., Harley V.R., Hudson P.J., Azad A.A.;
 RT "Sequence of the small double-stranded RNA genomic segment of
 RT Infectious bursal disease virus and its deduced 90-kDa product."
 RL Virology 163:240-242(1988).
 CC - FUNCTION: THIS IS THE PRESUMPTIVE VIRION-ASSOCIATED RNA-DEPENDENT
 CC RNA POLYMERASE. IT MAY ALSO CONTAIN GUANYLYL AND METHYL
 CC TRANSFERASE ACTIVITIES.
 CC - CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
 CC (RNA)(N).
 CC - PTM: MAY EXIST IN MULTIPLE PHOSPHORYLATED FORMS.
 CC - SIMILARITY: BETWEEN THE CENTRAL REGIONS OF IBNV AND IBDV VPI.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M19336; AAA89177.1; -
 CC PIR: A28649; RRSXB.
 DR Transferase: RNA-directed RNA polymerase; Phosphorylation;
 KW GTP-binding.
 KW 411 422 RDRP CONSENSUS ELEMENT-1.
 FT DOMAIN 483 498 RDRP CONSENSUS ELEMENT-2.
 FT DOMAIN 527 540 RDRP CONSENSUS ELEMENT-3.
 FT NP_BIND 258 265 GTP (POTENTIAL).
 FT SEQUENCE 878 AA: 97274 MM: 6B67C1E0AC193D1F CRC64;
 SQ
 Query Match 2.8%; Score 147.5; DB 1; Length 878;
 Best Local Similarity 20.2%; Pred. No. 0.026;
 Matches 157; Conservative 109; Mismatches 272; Indels 241; Gaps 39;

QY 17 PPRADSEFOALDIMEPEKPLPT-----AFIVDTSEAL-----LLPD 56
 DB 22 PTAGQDVEELLTPKVVWPPEDPLASPSRLAKFLRENGIKILQPSLPENEEYETDILPD 81
 QY 57 WLKLMIRSEVL-----RLVDAALQDLEPOQ-----LLLVQSF 90
 DB 82 LAMWRQTEGAVALKRTLSLHMGPRVLKPVLLNSPREGQKACVPRHCTQADIVLFEO-- 139
 QY 91 GIPVSSMSKLLQFLDQAVAHDPQTLQONIMKNK-----MAHLYEVQHERASG 139
 DB 140 -VPEAT-----ESLKEVETLLTQINIRDKAAGSGTYGQATRLVAMEVATGRPNK 189
 QY 140 -----GOTFS--LLASLPPRDSREAPKPKSPQPIGGRIRVGTOLRVLPEDDL 191
 DB 190 DPLKLGTFESIAQLDITLP-----VGPEEDDKPVVPL--TRVPSRLVLY--TGDV 238
 QY 192 AGMFLQIFPLSPDRMOSPPR-VALALQALQOELARVAVQGSPEVQGITVRLQALAT 250
 DB 239 DGDG-EVEDLPLKINLKSSSGLPVYVGRTKGETIGEMIA-----ISNQLRELSA 286

QY 173 -GGRIRVGTOLRVLPEDDLAGMELQIFPLSPDRMSSSPRPALALQ----- 221
 DB 3224 LQGMQKRL--MKGED--GGHSL-----SSVR-----TLQOQRRRLERELE 3262
 QY 222 ALGQELARV-----QGSPEVGTIVRYLQALATL--ISSPHGALVMSHNSHFLA- 271
 DB 3263 AMEREVARLQTEACRLQGLHPAARGLAKVOEAMATLQAKQERGQWMLAQQAQGHAFGR 3322
 QY 272 CPLL-----ROLQYQRCVPODGFSSFLKLVLLQMLQMDSCVEGCEPLRAQLRMLAS 325
 DB 3223 COELLANAQEQELASSEELAEVDVAGAFQQLQGHHELGQETRECRLOQODLRQEQQLVD 3382
 QY 326 -----QASAGRRLSDVNRGLRLALAEALAFROD-----LEVSSTVAVIATLTERSGQC 373
 DB 3383 NSHFMASVEFTECQLQELBESRLQELBEAMALRWQRCAESGCLKRLRQLQAEVLAQWGL 3442
 QY 374 SVEPDL--ISKVLQGLIEVSPHLELLTFNSATADASPFPACKPVVVSSLLQEE 430
 DB 3443 LKRDYGHSSDV--ELLHRHODEKLKLA-----QEEKFAQOMQKTEMEQOELLLOPQ 3493
 QY 431 EPLAGKRGADGSLAEVRLGPS--SGLLVDMLEMLDEPVVSSCPDLQLRLFSRRKKG 488
 DB 3494 EL-----KPRAGSSLTFSQMPRSGHQIGAQDLAETRDQDAKKGPTMGSLFQKHLPLG 3549
 QY 489 QAQVPSFRPYLLLTFTQSSWPTLHQCIRVLGKSRQGRFPDPSASLDPLMACIHVPRIMO 548
 DB 3550 GRQ-PS-----SSWDS--CGTLQGSLSLFLDERMAAEKV-ASIALDLDTG 3593
 QY 549 GRDORTQK--RREELVLRV--QGPPELLSVLELLAEATRSOD-----GDTAACSL--I 597
 DB 3594 ARCLRLGRHGRKHTFSRLTSGAEI-----LFAAPSEQAEASWRAUGSTAALSLPKL 3648
 QY 598 QARPLPLSCCCGDD 612
 DB 3649 KAKPVSLNECTTKD 3663

RESULT 5
 ID T2D3 HUMAN STANDARD; PRT: 1083 AA.
 AC 000268; Q99721; Q9BX42; Q9BR40;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transcription initiation factor TFIIID 135 kDa subunit (TAFI1-135)
 GN TAFI135 (TAFII-130) (TAFII130).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97336072; PubMed=9192867;
 RA Mengus G., May M., Carre L., Chambon P., Davidson I.;
 RT "Human TAF(II)135 potentiates transcriptional activation by the AF-2s
 of the retinoic acid, vitamin D3, and thyroid hormone receptors in
 RT mammalian cells.";
 RL Genes Dev. 11:1381-1395(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leivasalho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurtry A.A.,
 RA Milne S.A., Misty D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.T., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [3]
 RP SEQUENCE OF 105-1083 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=97098442; PubMed=8942982;
 RA Tanese N., Saluja D., Vassallo M.F., Chen J.-L., Admon A.;
 RT "Molecular cloning and analysis of two subunits of the human TFIIID
 RT complex: hTAFII130 and hTAFII100.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13611-13616(1996).
 CC -I- FUNCTION: MAKES PART OF TFIIID IS A MULTIMERIC PROTEIN COMPLEX THAT
 CC PLAYS A CENTRAL ROLE IN MEDIATING PROMOTER RESPONSES TO VARIOUS
 CC ACTIVATORS AND REPRESSORS. POTENTIATES TRANSCRIPTIONAL ACTIVATION
 CC BY THE AF-2S OF THE RETINOIC ACID, VITAMIN D3 AND THYROID HORMONE.
 CC -I- SUBUNIT: TFIIID IS COMPOSED OF TAFI BINDING PROTEIN (TBP) AND A
 CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFs).
 CC -I- SUBCELLULAR LOCATION: Nuclear.
 CC -I- SIMILARITY: BELONGS TO THE TAF2C FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is thus stated. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: Y11354; CAA72189.1; -;
 DR EMBL: AL137077; CAC36006.1; -;
 DR EMBL: AL109911; CAC23212.2; -;
 DR EMBL: U75308; AAC50901.1; -;
 DR TRANSFAC: T02328; -;
 DR Gene: HGNC:11537; TAF4.
 DR MIM: 601796; -;
 DR InterPro: IPR003894; TAF_hom.
 DR SMART: SMO0549; TAFH. 1.
 KW Transcription regulation; Nuclear protein.
 FT DOMAIN 39 42
 FT POLY-HIS.
 FT DOMAIN 52 57
 FT POLY-ALA.
 FT DOMAIN 98 101
 FT POLY-GLY.
 FT DOMAIN 142 148
 FT POLY-ALA.
 FT DOMAIN 268 275
 FT POLY-PRO.
 FT DOMAIN 331 337
 FT POLY-ALA.
 FT DOMAIN 680 683
 FT POLY-PRO.
 FT DOMAIN 808 813
 FT POLY-ALA.
 FT DOMAIN 828 831
 FT POLY-ASP.
 FT DOMAIN 105 117
 FT PGPPSPRRPLVPA -> GRGLIQGRGREGS
 FT (IN REF. 3).
 FT CONFLICT 136 136 A -> S (IN REF. 2).
 FT CONFLICT 185 185 G -> GCG (IN REF. 2).
 FT CONFLICT 233 264 MISSING (IN REF. 3).
 FT CONFLICT 293 293 P -> L (IN REF. 3).
 SO SEQUENCE 1083 AA; 109943 MW; A6453827572A0752 GRC64;
 Query March 2.5%; Score 133; DB 1; Length 1083;
 Best Local Similarity 20.8%; Pred. No. 0.35; Indels 202; Gaps 28;
 Matches 135; Conservative -----LTAALPPRRDSTEARPKKSSP-----EQPIGGIRIRV 178

```

Db 400 KGAGAVT-QSLSRPTATTGIRATLTPTVLAPRLPQPQNPNTNIONFGLPGMVLVRS 458
Qy 179 GTQLRVLGEEDDLAGNFIQIF-----PLSPDPWSSSPVALLAQALQGE-LARVVO 232
Db 459 ENGQGLMIPOQALAQMOQAHNAPQPTMARPTPTSAF-PVOISTYQAGCTPIANQVT 517
Qy 233 GSEPEVPGIT--VVALATATLSSPH-----GALVMSHSHFLACPLRLQOLCOYOR 283
Db 518 PTTIIRKQVSOAQCTVOPSAFLQRSFGVQPOLVVGAA--OTASIGTATAVOTGTPQR 572
Qy 284 CVPQDGFSSL-----FLKVLLOMQWIDSPVEGGLRAQRLMASQASAG 330
Db 573 TPGATTSSATETMENKCKNFSLTIK-----LAASSKQSTETANVKEVLQNLIDG 628
Qy 331 R-RLSDVRGGLR-----LAEALAPRODLE----- 354
Db 629 KTAEDFTSLRYELNASSPQRYLVPLKRLPALRQULTPSAFLIQSOQOQPPPTSQAT 688
Qy 355 -----VVSIVR-----AVIATLRSGEQSV-----EPDLISKVLQG----- 386
Db 689 TALTAVLSSSVORTAGTAAVTASALQPVLSLTQPTQVGVKQCPPTPLVIQPPKPG 748
Qy 387 -----LIEVRSPHLELTAFESATADAASFPAKCPVYVVSLLQEEEP 432
Db 749 ALIRPQVTLTQTPVVALRQPHRIMLTTPQOIQLNPLQVPPVKKPVALCTALSAVSA 808
Qy 433 LAGG-----KPGA-----DGSLEAVRLGSSGLVWLEMLDPEVSSCPD- 474
Db 809 QAAAGAKNKLEPGGSGFRDDDDINDVASMAGVNLSEASRIATNSELVGTILTRSCKE 868
Qy 475 -----LQRL-L-FSRKRGKGAQVPSFPRYLLTFTHOSSMPTLHQCIRVLGK----- 522
Db 869 TELLAPLQRLRIETLICKKH-----ITELHPDVSVYSNAT-----QORLONLVEKISET 918
Qy 523 -----SREORFDSASLDFLMACIHPVIMOGDORTPOKRELEV-----L 564
Db 919 AOKNFYSKDDDRYEQASDVRAOLKFF-----EQDDQIKRQKDEDERITILRAKRSR 972
Qy 565 RVQGPPELISL-----VELILAE-AETRSQDGDITACSLIQARPLLLSC 607
Db 973 RQDPEQLRLKQAKEMQOELAQMRORDANLTALAIIGRKKRKYDC 1020

```

Result 6

PCN2_HUMAN STANDARD; PRT; 3321 AA.

AC 095613; 043152;

DT 30-MAY-2000 (Rel. 39, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Pericentrin 2 (Kendrin).

GN PCN2 OR KIAA0402.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Li O., Joshi H.C.;

RT "Kendrin, an integral component of centrosome."

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 1512-3321 FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=98116655; PubMed=9455477;

RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,

RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. VIII.

RT 78 new cDNA clones from brain which code for large proteins in

RL DNA Res. 4:307-313(1997).

RN [3]

```

RP REVISIONS, AND SEQUENCE OF 1192-3321 FROM N.A.
RA Ohara O.
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 787-1522 FROM N.A.
RC TISSUE=Trachea;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsuyama Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Katsuyama S., Satoh N., Matsunawa H., Takahashi E., Teraoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Ota T.,
RA Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hiro T., Saito K.,
RA Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y.,
RA Sekine M., Kikuchi H., Kanda K., Wagatsuma M., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahori K., Masuho Y., Nagai K., Iisaga T.,
RT "MEDO human cDNA sequencing project."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CENTROSOMAL PROTEIN.
CC -1- SIMILARITY: STRONG, TO MOUSE PERICENTRIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).


DR EMBL: U52962; AAD10838.1; -



DR EMBL: AB007862; BAA23698.2; ALT_INIT.



DR EMBL: AK093923; BAC04252.1; ALT_INIT.



DR Genew: HGNC:16068; PCNT2.



DR MIM: 605925; -



KW Coiled coil.



FT DOMAIN 258 553 COILED COIL (POTENTIAL).



FT 675 835 COILED COIL (POTENTIAL).



FT 999 1135 COILED COIL (POTENTIAL).



FT DOMAIN 1288 1938 COILED COIL (POTENTIAL).



FT 2053 2071 COILED COIL (POTENTIAL).



FT DOMAIN 2525 3075 COILED COIL (POTENTIAL).



FT 819 819 A -> G (IN REF. 4).



FT 879 879 A -> T (IN REF. 4).



FT 891 891 Q -> Q (IN REF. 4).



FT 956 956 T -> S (IN REF. 4).



FT 1013 1013 K -> E (IN REF. 4).



FT 1025 1025 T -> A (IN REF. 4).



FT 1276 1276 I -> L (IN REF. 1).



FT 1306 1306 K -> T (IN REF. 1).



FT 1523 1523 H -> Q (IN REF. 2).



FT 2177 2177 R -> M (IN REF. 2).



FT 2538 2538 T -> A (IN REF. 2).



FT 2828 2906 MISSING (IN REF. 2).



FT 3125 3125 S -> Y (IN REF. 2).



FT 3289 3289 G -> E (IN REF. 2).



FT 3309 3321 YVQILQNRNPNR -> VLPDSTSKSKSCHPMIKQ (IN REF. 2).



SO SEQUENCE 3321 AA; 376355 MW; AOB693FACADA151 CRC64;



Query Match 2.4%; Score 127; DB 1; Length 3321;



Best Local Similarity 20.4%; Pred. No. 4.5; Mismatches 284; Gaps 39;



Matches 187; Conservative 117;



Qy 17 PPRADSEFOALDIWEPEKPLPTAFVDTSEALLPDMKLRLMIRSEVRLVDAALQ 76



Db 2411 PPRKED-EIQQ-LSINGKQGEVPTA-----CPDW-----RGGLLVQGEAFE 2451



Qy 77 -DLEPQQLLFVSGFIPVSSMSKLLQFLDQAVAHDPQTEQINIMDKNVAHL----- 128



Db 2452 KEDEMOGVLELPRLSGSDLGSHSLERLEKIR-----EQDDLEKSLERLRLPDRSS 2505



Qy 129 -----VEVQHERGASGGOTFHSLLTALPPRDRSTEARPKRSSPEQPIGGRR 177


```

```

Db 2506 LTSEIOLRAOLRMTHLNOEKLHRTALT-----STBA-----RGSQOH-----Q 2548
Oy 178 VGTOLRVGLPEDDLAGMFLQIFPLSPDRWSSSPRYALALQALG--QELARVYQSGP 235
Db 2549 LRRVYELLA-----YKVEQEKCIAGDLOKTLSEOEKANSVOKLL 2588
Oy 236 EYPGTVTVYLOALATLSSPHGALVMSHRSHTLAPLRLQCYORCPVDTGSSIF 295
Db 2589 AAEQVYVNDLKS-----DLCEKRSKSEOLS-----RSLCEVQGVLDLRKSLSSK 2633
Oy 236 LKVLLOMLQWLDSPVEGGPLRA-----OLRMLASQASAGRRLSDVRGGL-----LRL 343
Db 2634 ENELAAALOELESEOGKGAHQSOLEEEOQLRHIORESQAKALEELRASLETORAOSSRL 2693
Oy 344 AEAALFRQ-----DLEVSSYTRAVIATIRSGEQCEVEDLISKVLOGLIEVRSPL 395
Db 2694 CVALKHEQAKNDLOKEKLEIHSRCEALIAQERS--OLSELQDLAAEKSRTELESLR 2752
Oy 336 BELTAFFSATADASPFPACKPVVYVSSLLLOEEPELAGKPKGADGSLAVRLCPSSG 455
Db 2733 EHLTLBOQLSQRTQACVHDQ-----AHNALLQ-----KLKEKS 2788
Oy 436 LLYVWLEMLDPEVSSCPDLQRLLFSSRRKGGQAQVPSFRPYLLTLFTHOSSWPTLLHC 515
Db 2789 RYVDLOAMLE-----KVQOQALHSQOQLEAEAO-----KHC 2819
Oy 516 IYVLLKSKSEQRFPDSASIDFLMACIHYRIMQGRDQRTPOKRRRELYRVQGPELTSIV 575
Db 2820 -----EALRRKEVSATLKTSTVEALHT--OKRELKSLERERKRPAMLO-----A 2862
Oy 576 ELILAETRSQDGTAAACSLIOARPLLLSCCGDSEVRYTEHLSCIOQMGDSVLG 635
Db 2863 ELEQSHPRKKEBEGKKAARSAEAR-----QSPAADQMKR--W 2899
Oy 636 RRCRDLILQLYORP-----ELRVYVPEVLLHSEGAASSVCKLDGLIH 679
Db 2900 QKDKERLELELQORDLHKIKOLOTVYRDLSEKDEVPGRSLHLSARRAASDADHL-- 2957
Oy 680 RRTTLADTSDSRALENRGADASMKRLAVAHPLLLRHLRMIYALLHGRHNLNQER 739
Db 2938 -----REOQRELE-----AMORRLSAA-----RLTSFTSQAVDRT-VNDWTSS 2996
Oy 740 QONHLSCFLHYGLL--ELLQPHVRSSEHOGALMDCLSFIRLLNYRKSSEHLAFLNK 797
Db 2997 NKKAVMSLHLELELKSLSR--TSQKKMAELQPFVYLL--KDNVSLTALST 3050
Oy 798 FVQFIHKYITTYAPPAISFLQKHADPLHDLSTFNSDLYMLKSLLAGSLPSRDRDTDGL 857
Db 3051 VTQ--EKLELSRAVSKLEKILKH-----HLOKGS--PSRSERSAMPK 3089
Oy 858 DEEGEES-----SAGSLP 871
Db 3090 DETAFQSSLRPDPGRLP 3107

RESULT 7
SBCC_PSEAE
ID SBCC_PSEAE STANDARD: PRT: 1211 AA.
AC Q9HMB8:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nuclease sbccd subunit C.
GN SBCC OR PA4282.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
MEDLINE=20437337; PubMed=10984043;

```

```

RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garner R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lathig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -!- FUNCTION: Sbbc cleaves DNA hairpin structures. These structures
CC can inhibit DNA replication and are intermediates in certain DNA
CC recombination reactions. The complex acts as a 3'->5' double
CC strand exonuclease that can open hairpins. It also has a 5'
CC single-strand endonuclease activity (By similarity).
CC -!- SUBUNIT: Heterodimer of sbcc and sbcd (by similarity).
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SBCC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC DR EMBL: AE004844; AAC07670.1; -.
CC Interpro: IPR003439; ABC_transport.
CC
CC Hydrolyase; Nuclease; Exonuclease; Endonuclease; DNA replication;
CC NP_BIND 37 44 ATP (POTENTIAL).
CC FT DOMAIN 115 156 COILED COIL (POTENTIAL).
CC FT DOMAIN 197 407 COILED COIL (POTENTIAL).
CC FT DOMAIN 497 584 COILED COIL (POTENTIAL).
CC FT DOMAIN 628 1048 COILED COIL (POTENTIAL).
CC FT SEQUENCE 1211 AA; 138959 MW; FC8DCA65C0B8E2 CRC64;
SQ
Query Match 2.4%; Score 125.5; DB 1; Length 1211;
Best Local Similarity 21.0%; Pred. No. 1.3;
Matches 232; Conservative 118; Mismatches 426; Indels 329; Gaps 49;
Oy 24 EFOALDI-----WPEEKPLPTAFVDTSEELLPMLKLMIRSEVRLVDALQDL 78
Db 251 EEOALQRLGEOQOWFTEER-----LLQSCENH-----QGLAEARQAWDALATER 296
Oy 79 EPOQLLFYQSGFIPYSSSKLQFLQDAVANDPOTLEONIMDKNMAHLYEVOHERGAS 138
Db 297 ETLQWLERL-----APYKGLIERLKQLEQLRHSSEQOR-----RTQQA 338
Oy 139 GGQTFHSLTASLPPRRDSTEAPKPSPEOPRIGGRINVTOL-----RVLGPEDD 190
Db 339 GTERLQGLQRLQEARERQADNHLQADP-----LREARQLSEARRELTTLAEBOE 393
Oy 191 LAGMFLQIFPLSPDRWSSSPRYALALQALGOLARVYQSGPVPVGTIVYVLOALAT 250
Db 394 L-----HROSQNRHAQOSDARQOLMDQQRHVAEOQLQ--ALRDSQALAA 438
Oy 251 LTSS--PHGALVMSHRSHTLAPLRLQCYORCPVDTGSSIFLVL-----LOML 303
Db 439 LGDAWTHOGQALATFYQR-----QRLAEQAOQLPELEKSLAHNGEFLRL 484
Oy 304 --OWLDSPEVEGGPLRAOLRLMLASQASAGR-----LSVGRGLRLAEALAFRO 351
Db 485 QAWMTALHSEPDILARLVELLRQDSTLERQDALHKEWQVLDQRAAGLARRGEL---- 540
Oy 352 DLEVSSYTRAVIATIRSGEQCEVEDLISKVLOGLIEV-----RSHPLELLAFRS 404
Db 541 DGRWEOE--QALDLKRGSSQCAEEYKAAQALQVTRILLQORRLARSASVEQLRAGLYD 599
Oy 405 ATADAASPPACKPVVYVSSLLLOEEPELAGKPKGADGSLAVRLCPSSGLVDYLEML 464
Db 600 GER-----CPVCG-----SQENP-----YHSEBQLLAALGEHD 627
Oy 465 DPEVSSCPDLQ-----LRLIFS--RRKGGQAQVPSFRPYLLTLFTHOSSWPTLLH 513

```

DB 628 DOQVRAEOSTLERLQTLVGLREGYSSQREBLNOSROEOELTQGLALROLDOH-TLP 686
 QY 514 QCIIRVLLGKREQREDFPSASLDFIMACIHPRIHQGRDQTP-OKRAEELVRYOGELI 572
 DB 687 EELRL- - - - -OPSADLEML- - - - -AOLDDLAGQKOCORPDRILAROTQOL 732
 QY 573 SLVELLAEAEYSQDDTACSLIQARLPULLSCCGDESVAKV-TEHLSGCIQOM- 629
 DB 733 - - - - -QOEIRAEETLQORQALTEBQ-RYEHLQOOVEEDSQOLRPLSDEHW- - - - -QRMQ 784
 QY 630 - - - - -GDSYLGRCRDLLOLYLOR-PEIRVPE- - - - -VLHSEG 665
 DB 785 ADPLRTFOALGESIEQRQOARLQOEORLQKORCDSSWQKQSDQORNEARQAE 844
 QY 666 AASSVCKLGLDIHRTITLADTSD-SRALENRCADSMACRKLAVAHPLHLMPIA 724
 DB 845 RAQAEELAEINGRLGALHGNACADQWOLSLFHAAQAAQAVETIQA-PLDSLREOL- 900
 QY 725 ALLHGRTHLNFQEFROQNH- - - - -SCFLHVLGL- - - - -L 754
 DB 901 RLAEALHLLQOORQORQOQDQWREBODNLDSRLDLGLSEBOATQWREOL 960
 QY 755 ELLOPHFR- - - - -SEHOGALMCLLSFIR- - - - -LLNKRKSRHLAATINKFVQ 800
 DB 961 QRIQOEETRQOTLEAEQAO- - - - -LQHRORPETDREALEDNROQRERLAEQAYLE 1016
 QY 801 FHKYITYNAPAAISFLQKADPLHDLSPNSDLYMLKSLAGLSLPSRDRTRGDEE 860
 DB 1017 - - - - -TYS- - - - -QLOADNORREOQALLAEERARERGRRLNE- 1053
 QY 861 GEEBSAGSLPLYSVSLFTPLTAEMAPYMKRISRGOTVEDLEVLSDID-EMSR- 915
 DB 1054 - - - - -LIGSSGDKFRRIQOYNL-DLVQHSNVOLQOLARRYRL 1092
 QY 916 - - - - -RPELISFSTNLQRLMSAECCNRLAFSLARSONSP-SIAA 958
 DB 1093 QRGSEGLLVDTKEMDELRSVYS- - - - -LSGGEFLISLALGLASWASKLRIES 1146
 QY 959 AFLPTFMYCLGSDFEVYOTALRNL 983
 DB 1147 LFIDEGFGLDPESLQAMDALDNL 1171

RESULT 8
 ID FACA_HUMAN STANDARD: PRT; 1455 AA.
 AC 015360; 092497; 075266; 09Y6M2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fanconi anemia group A protein (FACA protein).
 GN FANCA OR FAA.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97312685; PubMed=9169126;
 RA Ianzano L., D'Apolito M., Centra M., Savino M., Levran O.,
 RA Auerbach A.D., Gleason-Jansen A.M., Doggett N.A., Pronk J.C.,
 RA Callen F.C., Zelante L., Savoia A., Whitmore S.A., Apostolou S.,
 RT "The genomic organization of the Fanconi anemia group A (FAA) gene";
 RL Genomics 41:309-314(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphoblast;
 RX MEDLINE=97051928; PubMed=8896563;
 RA Lo Ten Foe J.R., Roolmans M.A., Bosnoyan-Collins L., Alon N.,
 RA Wijket M., Parker L., Lightfoot J., Carreau M., Callen D.F.,
 RA Savoia A., Cheng N.C., van Berkel C.G.M., Strunk M.H.P., Gilie J.J.P.,

RA Pals G., Kruyt F.A.E., Pronk J.C., Arwert F., Buchwald M., Joenje H.;
 RT "Expression cloning of a cDNA for the major Fanconi anemia gene,
 RT FAA";
 RL Nat. Genet. 14:320-323(1996).
 RN [3]
 RP SEQUENCE OF 175-834 FROM N.A.
 RA Rieke D.O., Bruce D., Mundt M., Doggett N., Munk C., Saunders E.,
 RA Robinson D., Jones M., Buckingham J., Chasteen L., Thompson S.,
 RA Goodwin L., Bryant J., Tesmer J., Melnick L., Longmire J., White S.,
 RA Ung S., Tatum O., Campbell C., Fawcett J., Maltbie M., Deaven L.;
 RT "Sequencing of human Fanconi anemia complementation group A gene
 RT genomic region";
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 491-542 FROM N.A.
 RX MEDLINE=98375690; PubMed=9711872;
 RA Levran O., Doggett N.A., Auerbach A.D.;
 RT "Identification of Alu-mediated deletions in the Fanconi anemia gene
 RT FAA";
 RL Hum. Mutat. 12:145-152(1998).
 RN [5]
 RP SUBCELLULAR LOCATION, AND MUTAGENESIS.
 RX MEDLINE=98414611; PubMed=9742112;
 RA Naef D., Kupfer G.M., Suliman A., Lambert K., D'Andrea A.D.;
 RT "Functional activity of the Fanconi anemia protein FAA requires FAC
 RT binding and nuclear localization";
 RL Mol. Cell. Biol. 18:5952-5960(1998).
 RN [6]
 RP PHOSPHORYLATION.
 RX MEDLINE=99007271; PubMed=9789045;
 RA Yamashita T., Kupfer G.M., Naef D., Suliman A., Joenje H., Asano S.,
 RA D'Andrea A.D.;
 RT "The Fanconi anemia pathway requires FAA phosphorylation and FAA/FAC
 RT nuclear accumulation";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:13085-13090(1998).
 RN [7]
 RP VARIANTS FA, AND VARIANTS
 RX MEDLINE=98058767; PubMed=9371798;
 RA Levran O., Erlich T., Magdalena N., Gregory J.J., Batish S.D.,
 RA Verlander P.C., Auerbach A.D.;
 RT "Sequence variation in the Fanconi anemia gene FAA";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:13051-13056(1997).
 RN [8]
 RP VARIANTS FA PRO-1110 AND GLY-1117.
 RX MEDLINE=99225041; PubMed=10210316;
 RA Kupfer G., Naef D., Garcia-Higuera I., Wasik J., Cheng A.,
 RA Yamashita T., Tipping A., Morgan N., Mathew C.G., D'Andrea A.D.;
 RT "A patient-derived mutant form of the Fanconi anemia protein, FANCA,
 RT is defective in nuclear accumulation";
 RL Exp. Hematol. 27:587-593(1999).
 CC -1- FUNCTION: DNA REPAIR PROTEIN THAT MAY OPERATE IN A POSTREPLICATION
 CC REPAIR OR A CELL CYCLE CHECKPOINT FUNCTION. MAY BE IMPLICATED IN
 CC INTERSTRAND DNA CROSS-LINK REPAIR AND IN THE MAINTENANCE OF NORMAL
 CC CHROMOSOME STABILITY.
 CC -1- SUBUNIT: BELONGS TO A MULTISUBUNIT COMPLEX COMPOSED OF FANCA,
 CC FANCB AND FANCG PROTEINS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (MAJOR) AND CYTOPLASMIC (MINOR).
 CC -1- PTM: PHOSPHORYLATED PRIMARILY ON SERINE RESIDUES. PHOSPHORYLATION
 CC IS REQUIRED FOR THE FORMATION OF THE NUCLEAR COMPLEX. NOT
 CC PHOSPHORYLATED IN CELLS DERIVED FROM GROUPS A, B, C, E, F, G, AND
 CC H.
 CC -1- DISEASE: DEFECTS IN FANCA ARE THE CAUSE OF ONE OF THE 8
 CC COMPLEMENTATION GROUPS OF FANCONI ANEMIA (FA), AN AUTOSOMAL
 CC RECESSIVE DISORDER CHARACTERIZED BY PROGRESSIVE PANCYTOPENIA, A
 CC DIVERSE ASSORTMENT OF CONGENITAL MALFORMATIONS, AND A
 CC PREDISPOSITION TO THE DEVELOPMENT OF MALIGNANCIES. AT THE CELLULAR
 CC LEVEL IT IS ASSOCIATED WITH HYPERSENSITIVITY TO DNA-DAMAGING
 CC AGENTS, CHROMOSOMAL INSTABILITY (INCREASED CHROMOSOME BREAKAGE),
 CC AND DEFECTIVE DNA REPAIR. FANCA DEFECTS ACCOUNT FOR 65 % OF FA.
 CC -1- DATABASE: NAME=Fanconi Anemia Mutation database;
 CC WWW="http://www.rockefeller.edu/fanconi/mutate/";
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Hematol.;
 CC WWW="http://www.infobiogen.fr/services/chronocancer/genes/FA1D102.html".

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; 283067; CAB05445.1; JOINED.
 DR EMBL; 283068; CAB05445.1; JOINED.
 DR EMBL; 283069; CAB05445.1; JOINED.
 DR EMBL; 283070; CAB05445.1; JOINED.
 DR EMBL; 283071; CAB05445.1; JOINED.
 DR EMBL; 283072; CAB05445.1; JOINED.
 DR EMBL; 283073; CAB05445.1; JOINED.
 DR EMBL; 283074; CAB05445.1; JOINED.
 DR EMBL; 283075; CAB05445.1; JOINED.
 DR EMBL; 283076; CAB05445.1; JOINED.
 DR EMBL; 283077; CAB05445.1; JOINED.
 DR EMBL; 283078; CAB05445.1; JOINED.
 DR EMBL; 283079; CAB05445.1; JOINED.
 DR EMBL; 283080; CAB05445.1; JOINED.
 DR EMBL; 283081; CAB05445.1; JOINED.
 DR EMBL; 283082; CAB05445.1; JOINED.
 DR EMBL; 283083; CAB05445.1; JOINED.
 DR EMBL; 283084; CAB05445.1; JOINED.
 DR EMBL; 283085; CAB05445.1; JOINED.
 DR EMBL; 283086; CAB05445.1; JOINED.
 DR EMBL; 283087; CAB05445.1; JOINED.
 DR EMBL; 283088; CAB05445.1; JOINED.
 DR EMBL; 283089; CAB05445.1; JOINED.
 DR EMBL; 283090; CAB05445.1; JOINED.
 DR EMBL; 283091; CAB05445.1; JOINED.
 DR EMBL; 283092; CAB05445.1; JOINED.
 DR EMBL; 283093; CAB05445.1; JOINED.
 DR EMBL; 283094; CAB05445.1; JOINED.
 DR EMBL; 283095; CAB05445.1; JOINED.
 DR EMBL; 283151; CAB05445.1; JOINED.
 DR EMBL; X99226; AAC67610.1; -.
 DR EMBL; AC005567; CAC33401.1; -.
 DR EMBL; AF054569; AAC28331.1; -.
 DR EMBL; HGNC:3582; FANCA.
 DR MIM; 227650; -.
 DR InterPro; IPR003516; Fanconia.
 DR Pfam; PF03511; Fanconia; 1.
 DR PRINTS; PR00826; FANCONIAGEN.
 DR DNA repair; Nuclear protein; Phosphorylation; Disease mutation;
 KM Polymorphism.
 FT DOMAIN 18 34
 FT VARIANT 6 6
 FT VARIANT 8 8
 FT VARIANT 181 181
 FT VARIANT 244 244
 FT VARIANT 252 252
 FT VARIANT 435 435
 FT VARIANT 492 492
 FT VARIANT 501 501
 FT VARIANT 739 739
 FT VARIANT 809 809
 FT VARIANT 817 817
 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 V -> D.
 /FTid-VAR_009637.
 N -> K (IN FA A); COULD BE A
 POLYMORPHISM).
 A -> V (IN FA A).
 /FTid-VAR_009639.
 L -> F (IN FA A).
 /FTid-VAR_009640.
 D -> G (IN FA A).
 /FTid-VAR_009641.
 R -> C (IN FA A).
 /FTid-VAR_009642.
 H -> R (IN FA A).
 /FTid-VAR_009643.
 S -> G.
 /FTid-VAR_009644.
 P -> L.
 /FTid-VAR_009645.
 G -> D.
 /FTid-VAR_009646.
 L -> P (IN FA A).
 /FTid-VAR_009647.

FT VARIANT 845 845 L -> P (IN FA A).
 FT /FTid-VAR_009648.
 FT VARIANT 1055 1055 R -> L (IN FA A).
 FT /FTid-VAR_009649.
 FT VARIANT 1110 1110 H -> P (IN FA A); LOSS OF FUNCTION).
 FT /FTid-VAR_009650.
 FT VARIANT 1117 1117 R -> G (IN FA A); LOSS OF FUNCTION).
 FT /FTid-VAR_009651.
 FT VARIANT 1128 1128 O -> E (IN FA A).
 FT /FTid-VAR_009652.
 FT VARIANT 1131 1131 T -> A (IN FA A).
 FT /FTid-VAR_009653.
 FT VARIANT 1263 1263 MISSING (IN FA A).
 FT /FTid-VAR_009654.
 FT VARIANT 1287 1287 V -> I.
 FT /FTid-VAR_009655.
 FT VARIANT 1302 1302 W -> R (IN FA A).
 FT /FTid-VAR_009656.
 FT VARIANT 1328 1328 T -> A.
 FT /FTid-VAR_009657.
 FT VARIANT 1417 1417 H -> D (IN FA A).
 FT /FTid-VAR_009658.
 FT CONFLICT 717 717 M -> I (IN REF. 2).
 SQ SEQUENCE 1455 AA; 162804 MW; DB43EBDE29DF5C0E CRC64;
 Query Match
 Best Local Similarity 2.4%; Score 123.5; DB 1; Length 1455;
 Matches 256; Conservative 133; Mismatches 447; Indels 369; Gaps 63;
 QY 7 HAWVILLTLPGRADSEFOALLDIFPEERPLPTA-----FLVDTSEAL-LLPMLKL 60
 DB 322 HTLTQITLSPVLKASDAVOMQREMSFARTHPILTSLYRLFLVLSAEELVGHLOEVLE 381
 QY 61 RMIR-SEVLKVDALQDL-EPQQL-----LPVQSF-GIPVSSSKLLQFLDQAVAND 111
 DB 382 QEVHWQVLSFEVSALVVCPEEAQQLLEDVVARLMAQAFESQCDLSMTVAFLLVROALEG 441
 QY 112 PQLTEONIMDKNMAHLVEYQHERGASGQ-----TFHSLTLTASLPDRDSTE 159
 DB 442 PSA-----FLSTADMFKASFGSTRGHGSKALVFLTFLELVEFESPRYLQVH 492
 QY 160 APRKSSPEQ-----PLGGGRIRVG--TQLRVGLPEDDLGAMFLQIFPLSPDRWOS 209
 DB 493 ILHPVLVPSKYSRLTDYISLAKTRLADLKVSLENNGLYEDL----- 534
 QY 210 SSPRPVALALQALGQELARVYQGSPEVEGITRVLOALATLISPHGALVMSHRSH 269
 DB 535 SSAGDITPEHSQAL-ODVERAIWFETHGNIPVTME--ASTIRRP-----IVSH 583
 QY 270 LACPLRLQCYQYRCVPODTGFSSFLKVL-----LQMLQWLDSPEVEGGLPRAQT 320
 DB 584 LPAALL-----PRVLPKVPDSNVAFIESLKRADKIPPLSYISYVQACASAEKPPEDAL 637
 QY 321 RMLASQASGRRLSDVRGGLRLAELARQDLEVVSTYRAVIATLR-----SGQCSY 375
 DB 638 GVAEENSASAEELGQTLTALGELRASMTDPSQDRVISAQVAVISELRAVLGHNEDDSV 697
 QY 376 EPLILSKVGLGLEVSAPHEE-----LTLAFSATAADASPP----- 414
 DB 698 E--ISKI--QSLNTPRIEPRHMAVDLLTSFCQNLMAASSVAPPEBQWAAFLVR 751
 QY 415 -ACKPVV--VVSGL--LLOEEPLAGCKPGADGSLAVARLGPSSGLVDWLEMLDPEVY 469
 DB 752 TMGGRVLPVNLVRLQCLLRHGQSLSA-PRVGLALALAVHLGSGRSAL-----PEVD 802
 QY 470 SSCPDLQRLLESRRKRGKGOAVSFRPYLLTLFTHOSSMPTLHQC--IRVLGKSRQD 526
 DB 803 VGPP-----APGAG-LPVLPALFSLTLCRTDLSLFFCLFCTAAISYSLCKFFSSQ 851
 QY 527 RDPSPAS-----LDPL-----MACIHP-----RTW 547
 DB 852 SRDTLCSCSPGLIKKFOPLMFLRFLSEAROPLSSEEDVASLSMRPLHLPSPADKORALSLW 911

QY 548 QGRDPTPKRR-----ELVAVQPELISLV-----ELLIAEETR 585
 Db 912 THFRFEVLEEDVHTLYQDMLHELEIQ--PEADALSDTERQDPHQAHEHFPRESSA 970
 QY 586 SQ-DGDTAAACSLIOARPLLLSCCGGDESVRYKTEHLSCGICQMGSVLC----- 635
 Db 971 GGCGGD-----LOACTIIVNALMDHQSSRSY-DH-----SENSDLVGGGRGNEDIT 1018
 QY 636 RRCRDLILQYLQRPRLVPEVYLLHSEGAASSVCKLDGLHRTLLADTSDSRALE 695
 Db 1019 SRLQEMVADELQO-DLIVPL-----GHTPSQEHFLPEIFRRRIQALV-----S 1061
 QY 696 NRGDASMACRKLAVAPRLLLRLPMIALHRTHLNQEFQOQHLSC-----FLHVL 751
 Db 1062 GWSVAASLQRPRLVPEVYLLHSEGAASSVCKLDGLHRTLLADTSDSRALE 1112
 QY 752 GLLELQPHVFRSE-----HOGAL-WDCLLSFIRLLNRYKSSRHAAAFINKVFQFIHK 804
 Db 1113 -----NSBMRNFCSHGALITODITAHFRGLNACLRSRPSL-----WDFILA 1157
 QY 805 YITYNAPAAISFLQKADPHLDSFDSNDLVMLKSLAGL-----SLPSRDRDRTD 854
 Db 1158 KCQTKCPULIT-----SALVWMPLEPVLLCRMRHRCQSPRLPELQKL- 1200
 QY 855 RGLDEEBEESAGSLPLVSVSLFTPLTAAMAPYMKRLSRGQ--TVEDILE--VUSDI 909
 Db 1201 -----QEGROFASD-----FLSPEASPAENPDMLSAALHFAIQOYREENIKROL 1246
 QY 910 DEMSRRPPELISF-FSTNLORLMSAECCRNLAFLSLALRSNOMSPS--IAAFLPTF- 964
 Db 1247 KKLDCEBELLVLFEPFSLKGLSS-----HITSNTDLPFAHFCAAILLECLE 1296
 QY 965 -----MYCLGSQDEEVVQVOTALRNLPE-----YALLQOEHA-AVLLHRAFL- 1003
 Db 1297 KRKISMLALFQLESDRLRLRLLRVAPDQTRLLPFAFSLSYFHEDAIAIREAFILHV 1356
 QY 1004 -VGMV 1007
 Db 1357 AVDMY 1361

RESULT 9
 SPEC_HUMAN
 ID SRCO_HUMAN STANDARD; PRT; 2564 AA.
 AC Q9H254; Q9HCD0; Q9H3G8; Q9H1K7; Q9H1K8; Q9H1I9;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Spectrin beta chain, brain 3 (Spectrin, non-erythroid beta chain 3)
 DE (Beta-IV spectrin).
 GN SPTBN4 OR SPTBN3 OR KIAA1642.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=21316449; PubMed=11294830;
 RA Tse W.T., Tang J., Jin O., Korsgren C., John K.M., Kung A.L.,
 RA Gwyn B., Peters L.L., Lux S.E.;
 RA "A new spectrin, beta-IV, has a major truncated isoform that
 RT associates with promyelocytic leukemia protein nuclear bodies and the
 RT nuclear matrix";
 RL J. Biol. Chem. 276:23974-23985(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 TO 4).
 RX MEDLINE=20539976; PubMed=11086001;
 RA Bergs S., Aguijaro D., Dirix R.J., Maksimova E., Stabach P.,
 RA Hermei J.-M., Zhang J.-P., Philbrick W., Slepnev V., Ott T.,
 RA Solimena M.;
 RT "BetaIV spectrin, a new spectrin localized at axon initial segments
 RT and nodes of ranvier in the central and peripheral nervous system";
 RL J. Cell Biol. 151:985-1002(2000).

RN [3]
 RP SEQUENCE OF 386-2382 FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RX MEDLINE=20450683; PubMed=10997877;
 RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes.
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which
 RT code for large proteins in vitro";
 RL DNA Res. 7:273-281(2000).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms: 1 (shown here), 2, 3 and 4; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Abundantly expressed in brain and pancreatic
 CC islets.
 CC -1- SIMILARITY: BELONGS TO THE SPECTRIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 18 SPECTRIN REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF311855; AAC42473.1; -
 DR EMBL: AF082075; AAC38874.1; -
 DR EMBL: AY004226; AAF93171.1; -
 DR EMBL: AY004226; AAF93172.1; -
 DR EMBL: AY004227; AAF93173.1; -
 DR EMBL: AB046862; BAB13468.1; -
 DR HSSP: Q01082; IBKR.
 DR Genew: HGNC:14896; SPTBN4.
 DR MIM: 606214; -
 DR InterPro: IPR001589; Actbind.actnin.
 DR InterPro: IPR001715; Calponin-like.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR002017; Spectrin.
 DR InterPro: IPR001605; Spectrin_PH.
 DR Pfam: PF00169; PH; 3.
 DR Pfam: PF00307; CH; 2.
 DR Pfam: PF00435; spectrin; 56.
 DR PRINTS: PR00683; SPECTRINPH.
 DR SMART: SM00033; CH; 2.
 DR SMART: SM00150; SPEC; 16.
 DR PROSITE: PS00019; ACTININ_1; 1.
 DR PROSITE: PS00020; ACTININ_2; 1.
 DR PROSITE: PS50021; CH; 2.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 DR Cytokeleton; Membrane; Repeat; Actin-binding; Capping protein;
 KW Alternative splicing.
 FT DOMAIN 1 282
 FT DOMAIN 61 165
 FT DOMAIN 180 282
 FT REPEAT 309 354 CH 2.
 FT REPEAT 398 419 SPECTRIN 1.
 FT REPEAT 429 533 SPECTRIN 2.
 FT REPEAT 535 642 SPECTRIN 3.
 FT REPEAT 644 771 SPECTRIN 4.
 FT REPEAT 773 879 SPECTRIN 5.
 FT REPEAT 881 985 SPECTRIN 6.
 FT REPEAT 1019 1086 SPECTRIN 7.
 FT REPEAT 1088 1197 SPECTRIN 8.
 FT REPEAT 1199 1303 SPECTRIN 9.
 FT REPEAT 1305 1408 SPECTRIN 10.
 FT REPEAT 1410 1513 SPECTRIN 11.
 FT REPEAT 1515 1619 SPECTRIN 12.
 FT REPEAT 1621 1725 SPECTRIN 13.
 FT REPEAT 1727 1832 SPECTRIN 14.
 FT REPEAT 1834 1940 SPECTRIN 15.
 FT REPEAT 1942 2046 SPECTRIN 16.
 FT REPEAT 1942 2046 SPECTRIN 17.

```

FT REPEAT 2048 2107 SPECTRIN 18.
FT DOMAIN 2418 2527 PH.
FT VARSPLIC 1 1257 MISSING (IN ISOFORM 3).
FT VARSPLIC 1258 1286 AVOAAGELLROGNIYGEAOEAVTLLEK -> MHPYSCS
FT VARSPLIC 1287 1309 SAPSLGTPIPOITQOLEAHR (IN ISOFORM 3).
FT VARSPLIC 1287 1309 NOENOLRAOQWOKLDOLEOH -> CLTIHALLHPWE
FT VARSPLIC 1310 2564 PLYLRSS (IN ISOFORM 2).
FT VARSPLIC 2113 2154 MISSING (IN ISOFORM 2).
FT VARSPLIC 2113 2154 TEKIKAEOSKOPPTLGEDETELAANKAPILRPQY
FT VARSPLIC 2113 2154 E -> PREHDHNGVDOOPMOTHEKPSLPKAKKERTAR
FT VARSPLIC 2113 2154 RDGTCL (IN ISOFORM 4).
FT VARSPLIC 2155 2564 MISSING (IN ISOFORM 4).
FT CONFLICT 604 608 MISSING (IN REF. 2).
FT CONFLICT 714 714 L -> S (IN REF. 2).
FT CONFLICT 1189 1189 E -> K (IN REF. 2).
FT CONFLICT 1193 1193 E -> K (IN REF. 2).
FT CONFLICT 1331 1331 G -> S (IN REF. 1).
SQ SEQUENCE 2564 AA; 288982 MW; 52CDE7D1D6D01ECC CRC64;

```

Query Match 2.4%; Score 123.5; DB 1; Length 2564;
 Best Local Similarity 19.2%; Pred. No. 5.4;
 Matches 221; Conservative 152; Mismatches 451; Indels 325; Gaps 47;

```

QY 5 VYHAWVILTLCPRRAD-----SEFQALDIWPEEKPPLTAVLTSEALLIP 55
DB 940 VHTVOELVEGHPSSDEVSCODHLNRMNRIVELVEQRKEEMSAVLIVENHV----- 993
QY 56 DMLKLRIMSEY-----LRVDAALDLEPOQQLLF----- 86
DB 994 --LEVAEVAOYREKRRAVESAPRAGGALQWRISGLEALQLEPQALDEBALLAER 1051
QY 87 -----VQSGIPVSSMSKLFQDQAVHPDTLEONIMDKN-YMAHLEVEQHE 134
DB 1052 PPAQARLHQGAELGAEWGAELASANAQAGEAANAAGR-LQRFHLDLQFLQVLAQQA 1110
QY 135 RASGSGQTHSLTASLPRRST---EAPKPKSSPEQPIGQGRIRVGTQLRYLGEDDL 191
DB 1111 AGSGEPPRLNSLEADALLARHAALKEEVDREEDYARIVAASEALLADGAEELGP----- 1166
QY 192 AGMFLQIFLSPDPWQSS-----SPRPVALA-----LQOAL-----GOELARVQ 232
DB 1167 -GLADEMIRPHLEGNHKLGLMEARRELVQAHTIQLRLDRLQALVYLKRDQEA--LS 1223
QY 233 GSPEVPGITVRVLAATLTLSSPHGALVMSHRSHFLACPLRLQCY-----QRCPQ 287
DB 1224 GA-ELPGYVESYEAL-----KQHRDFLTMMELSQKMQVAVOAAEELRLQ 1268
QY 288 DTGFSLEFLKVLQMLQWIDSGVEGPIRAO--LRMLASQASAGRRISD----- 335
DB 1269 GNLYGEQAOEAVTRLLE-----KNQENQIRAOQWOKLHDOLELQHFELDCHELDGWIHE 1323
QY 336 -----VRGGL-----LR-----LAELAPRODLEVSSVYRAVITLRSGECCSVE 376
DB 1324 KMLMRDGRREDNHLKHLKHAQAFMAFLAONKEWLEIE-----REGQLMCE 1373
QY 377 -PDLISKVYLGLLEVRS--PHLEELLTAFFSATADAAPFPACKPVVVSSTL-----DL 427
DB 1374 KEPLAASVYKRLGELRQCAELES-----TQAKARQLEFASKDQQLVQSEFELDKLL 1427
QY 428 QDEEPLAGKPKADGSL--EAVRIGPSSGLLVDMLEMDPEVYVSSCPQLQ-----RL 480
DB 1428 HNESQLQDVDPGDLATVNSQLKQSMESQVEWYREVY-ELQACTALPLPEAPSKELY 1486
QY 481 FSRKRGKGOAVSPFRPYLLTLFTHQSSWPTLHOCIRVLYGSRREGRPSSSLDELMAC 540
DB 1487 GERQNAVERLRYLLP-----LQERRRLASKELHQVAHDDDELAWQ 1532
QY 541 IHVPRIMQ--GRDQTPQK-----RREELVLRVQGPPELISVLELLAEATRSQDG 589
DB 1533 ERLPLAMQTERNGQAVQOQIKKNOGLRE--IQAHGPRLEVEVLERGALASLRSPA 1589
QY 590 DTAAASLQARPLLLSCCGGDESVRYKTEHLSCIQMGWSVYGRRCRDLLQLYQOR 649

```

```

DB 1590 EA-----VRGLEQLQSANAGIREAERROVLDAAFO-----VEQYEDV 1630
QY 650 PELR--VPPPEVILHSEGAASSVCKLDDLIRFTLLADTSDSRATENRGADASACK 707
DB 1631 AEVEAWLQGEOLLMS-----EDKGDQESTLO- 1658
QY 708 LAVAPHLRLHPIALALHGRTHLNFQEFROQNHLSQFLVGLLELQCHVERSEHQ 767
DB 1659 -----LKKHLQLEQV-----ENYESTAQLSRQC-----RALLEGHDSQDISR 1701
QY 768 GALMDCLSFIRLLNRYKSSRHLLAFINKEVQFIHKYTYNAPPAISFLQKHADPLD 827
DB 1702 OSQVRLVYALRELGERR-----VALEQVYLYOLSRVSELE----- 1740
QY 828 SFDNSDLVNLKSLGLSLSPSRDRTDRGLDEGEPESS-----AGSLPVYSCLTP-- 880
DB 1741 -----HWIAEKEVAGSPPLGQDFEHVSYLQEFSSFASQMGKERLAAYNOMVDELI 1795
QY 881 -----LTAEMAPYMKRLSNGQTVEDLEVLSDIDEMSRRRPELTSFTNLQRLMSABE 936
DB 1796 ECGHTAAATMAEMKQGLN--EAMALLELMGTRQALLASRELKFFS-DARELQOIEE 1852
QY 937 CCRNLAFLSIALSMONSPSIAAFLPTNWCIGSODEFVYQALRNLPYA--LTCQ 991
DB 1853 KRRRLPRLTTPPEPRPSASSMORTIRAF-----EHDQLLVQVROLQGAQLRTVAG 1907
QY 992 EHAAYLHR 1000
DB 1908 EHAEAIASR 1916

```

RESULT 10
 MUTS.THECA
 ID MUTS.THECA STANDARD; PRT: 817 AA.

```

AC Q9ZIX6; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein muts.
GN MUTS.
OS Thermus caldophilus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_Taxid=272;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GK24;
RA Nashiru O., Park B.C., Ko J.H., Kim J.S., Koh S.K., Lee H.C.,
RA Kim C.H., Lee S.Y., Lee D.S.;
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: This protein is involved in the repair of mismatches in
CC DNA. It is possible that it carries out the mismatch recognition
CC step. This protein has a weak ATPase activity (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF007553; AAD01407.1; -
CC InterPro: IPR000432; MUTS_C.
CC InterPro: IPR002863; MUTS_N.
CC Pfam: PF00488; MUTS_C.1.
CC Pfam: PF01624; MUTS_N.1.
CC ProDom: PD001263; MUTS_C.1.
CC SMART: SM00534; MUTS_C.1.
CC SMART: SM00533; MUTS_G.1.
CC TIGRFAMs: TIGR01070; mutSL.1.
CC PROSITE: PS00486; DNA_MISMATCH_REPAIR_2.1.

```


Db 328 LPS-----LFNPNEDAFH-----EKTYSMD-----FVRRLRQCSQA 362
 QY 614 SVRKVTEH--LSGCIQOMGDSVLRRCRDLLOLYLO-----RELRVPEVILLHSG 665
 Db 363 SVKRLRAHAYHFNKKM-----LPVFOIRFRELASLEALDVL--EDA 408
 QY 666 AASVYCKLDGLIHR-FITLADTSDRALENGADAMACRKLAVAHLLRLHPLMA 724
 Db 409 PASPYCLLAS--HRTWSILRCWSEDFL-----PLVHR-----442
 QY 725 ALLHGTHTNPFQFROQNLSCPLAVLGLLELLQPHVRSERHOGALMDCLLSTRILLN 784
 Db 443 -----LMRULTL-----QILARY 454
 QY 785 RKSRLAIFINKFVQFIKITYTNAPAIISFLOKHADPLHLSFNSDLVLMKSLAGI 844
 Db 455 -----SYFVN--ELSLRPIISNEPKELK-----KPLVTS 482
 QY 845 SLPSSDDRTDRGLDEGESESSAGSLPLVVSILFPLTAEMAPYMKRLSGQTVEDLLE 904
 Db 483 KEPS--ITQGMTEDDGSGPSE--TKPVVSI-----SRQ---LVY 515
 QY 905 VLSDIDEMRRRPELISFSTNLQRLMSSAECCRLARSLARMSQNSPSTIAAF---L 961
 Db 516 VVADLDKLOEOLPELLEIKPKLEMI-----GPKNFSISALEDSQ 557
 QY 962 PTFMYCLSGSDFEYVQ 977
 Db 558 SFSFACVPSLSKIIQ 573

RESULT 12

EVPL_MOUSE STANDARD: PRT: 2035 AA.

AC Q9D952: 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Envoplakin (p210) (210 kDa cornified envelope precursor).
 GN EVPL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv.
 RA MEDLINE=20347896; PubMed=10747979;
 RA Maatta A., Ruhrberg C., Watt F.M.;
 RT "Structure and regulation of the envoplakin gene";
 RL J. Biol. Chem. 275:19857-19865(2000).
 RP [2]
 RP SEQUENCE OF 1860-2035 FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pancreas;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shigaawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
 RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinglich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,
 RA Hayashizaki Y.;

RT "functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -!- FUNCTION: COMPONENT OF THE CORNIFIED ENVELOPE OF KERATINOCYTES.
 CC MAY LINK THE CORNIFIED ENVELOPE TO DESMOSOMES AND INTERMEDIATE
 CC FILAMENTS.
 CC -!- SUBUNIT: MAY FORM A HOMODIMER OR A HETERODIMER WITH PPL.
 CC -!- SUBCELLULAR LOCATION: COLOCALIZED WITH DSP AT DESMOSOMES AND ALONG
 CC INTERMEDIATE FILAMENTS (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 7 PLECTRIN REPEATS.
 CC -!- SIMILARITY: CONTAINS 1 SPECTRIN REPEAT.
 CC -!- SIMILARITY: BELONGS TO THE PLEKIN OR CYTOLINKER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL: AJ309317; CAC38864.2; JOINED.
 CC DR EMBL: AJ319607; CAC38864.2; JOINED.
 CC DR EMBL: AJ319608; CAC38864.2; JOINED.
 CC DR EMBL: AJ319609; CAC38864.2; JOINED.
 CC DR EMBL: AJ319610; CAC38864.2; JOINED.
 CC DR EMBL: AJ319611; CAC38864.2; JOINED.
 CC DR EMBL: AJ319612; CAC38864.2; JOINED.
 CC DR EMBL: AJ319613; CAC38864.2; JOINED.
 CC DR EMBL: AK007353; BAB24980.1; ALT_INTT.
 CC MGD: MG1:107507; EVPL.
 CC InterPro: IPR001101; Plectrin_repeat.
 CC Pfam: PF00681; Plectrin; 3.
 CC SMART: SM00250; PLEC; 8.
 CC KIR: Keratinization; Repeat; Coiled coil; Cytoskeleton; Structural protein.
 CC KW DOMAIN 1 841
 CC FT DOMAIN 842 1674
 CC FT DOMAIN 1675 2035
 CC FT DOMAIN 12 28
 CC FT REPEAT 229 330
 CC FT DOMAIN 842 1664
 CC FT REPEAT 1186 1227
 CC FT REPEAT 1679 1714
 CC FT REPEAT 1819 1856
 CC FT REPEAT 1857 1894
 CC FT REPEAT 1895 1932
 CC FT REPEAT 1933 1970
 CC FT REPEAT 1971 2008
 CC FT CONFLICT 1861 1864
 CC FT SEQUENCE 2035 AA; 232317 MW; BB4A1D9CAED9641F CRC64;
 Query Match 2.3%; Score 122.5; DB 1; Length 2035;
 Best Local Similarity 19.9%; Pred. No. 4.5; Mismatches 361; Indels 373; Gaps 55;
 Matches 219; Conservative 145;
 QY 35 EEKPLPTA-----FLVDTSEALLLP-----DW--LKLIRSEVLRV 71
 Db 381 EEQGLAIARAVDDLDQRRQGEVAPLPQRRNPSSQPLHVSICWDSGEVQLRGECITLK 440
 QY 72 DAALDLEPOQLLFVOSFG-----IP-----VSSMKLLOFLDQAVAHDPQ 113
 Db 441 DNA-----DEYTWL--VQGGGGEFKSPACMLWIPADDPRAVCASTRL-----ATELD 486
 QY 114 TLEONIMDNVYMAHVEVOH-----EKGASG-----GOTFHSLLVASLPFRDS-----157
 Db 487 TLKQKLTENRKLKAAVEHLPOGQAPAGSAPADPQGGTLLSOMK-----QJGDGLGI 541
 QY 158 -----TEAPKRSSEPEPIGGRIR-----VGTOLRVGPEDDLAGMFLQJFPSPDR 206
 Db 542 ERQVLSMARSPLSQSSSLNDL-EGRIHSEGTAGRLQISGAKEANAQGCFAF-ISTKPT 599
 QY 207 WOSSSPRPV-----ALALQALGQELARVYG-----233
 Db 600 GSAAVQLPVVFSVKNRYNDVOSLCHLYGEEKAKAALGLEKQI-QEADRVIOGFEALALE 658

QY 293 SLFT----KVLLOMLQWLDSPGEGPLRAOLRMLASQASAGR----- 331
 DB 324 DLEIDSMALMAGLMDGHRGALGGPKALGSEGDADPGKALRGSEFPDGDVDD 383
 QY 332 -----RLSDVRGGLL-----RLAE-ALAFRODLEVVSVTRAVITLNSGECC----- 373
 DB 384 DRLVOEVRRLSATLGGGLQDHSRLSPGALPAPLKMERRKSERPEASLSEETACVE 443
 QY 374 SVEPDLISKVLOGLIEVRSPLHELLTAFSATADAAAPACKRVYVVSLLLOEPEPL 433
 DB 444 NVKSGTYSKDLG-----QPHSEPGAGAFGLQNMOPG-----SEEDSL 485
 QY 434 AGKRGADGSGLEAVRLGPSS-----GLVMDLEMDPEVSSCPDLQLRLFSRRKGQ 489
 DB 486 PAGQALGDLQ--LEVAKPESEEARGYIVTRDPLRPEEGQVADVARRLL----- 535
 QY 490 AOVPS 494
 DB 536 -QMP5 539

RESULT 14

GEM4_HUMAN STANDARD; PRT; 1058 AA.

AC P57678; Q9NZ57;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Component of gems 4 (Gemin4) (p97).
 GN GEMIN4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND MASS SPECTROMETRY.
 RX MEDLINE=20191885; PubMed=10725331;
 RA Charroux B., Pellizzoni L., Parkinson R.A., Yong J., Shevchenko A.,
 Mann M., Dreyfuss G.;
 RT "Gemin4: a novel component of the SMN complex that is found in both
 RT gems and nucleoli.";
 RL J. Cell Biol. 148:1177-1186(2000).
 CC -1- FUNCTION: THE SMN COMPLEX PLAYS AN ESSENTIAL ROLE IN SPliceosomal
 CC SNRNP ASSEMBLY IN THE CYTOPLASM, AND IS REQUIRED FOR PRE-MRNA
 CC SPLICING IN THE NUCLEUS. GEMIN4 COULD SERVE AS A COFACTOR OF
 CC GEMIN3.
 CC -1- SUBUNIT: FORMS A STABLE HETEROERIC COMPLEX WITH SURVIVAL OF MOTOR
 CC NEURON PROTEIN (SMN), GEMIN2 AND GEMIN3. INTERACTS DIRECTLY WITH
 CC GEMIN3 AND WITH SEVERAL SNRNP SM CORE PROTEINS, INCLUDING B/B',
 CC DI-D3, AND E.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR. LOCALIZED IN
 CC SUBNuclear STRUCTURES NEXT TO COILED BODIES, CALLED GEMS, WHICH
 CC ARE HIGHLY ENRICHED IN SPliceosomal SNRNPs AND IN THE NUCLEOLUS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF173856; AAF35283.1;
 DR Genbank: HGNC:15717; GEMIN4.
 DR MIM: 606969;
 KW mRNA processing; Spliceosome; Nuclear protein.
 FT DOMAIN 714 735 LEUCINE-ZIPPER (POTENTIAL).
 SQ SEQUENCE 1058 AA; 119989 MW; 5EE1FE2EE3538D6B CRC64;

Query Match 2.3%; Score 121.5; DB 1; Length 1058;
 Best Local Similarity 18.3%; Pred. No. 2.1;
 Matches 216; Conservative 149; Mismatches 359; Indels 457; Gaps 52;

QY 3 ILVHNAVILLTLPGRADSEFOALLDIWPEEKPPLTAFLVDTSEBALLPDWLKLRM 62
 DB 70 LIIIMAKV--LQPHVPTSDIETRMQEDLPSVGMIIPTI-----NHTLIFELKSTLE 120
 QY 63 IRSEVLRIVDAALDLEPQQLLEVQSGFIPVSSKSLLOFLD--QAVAH-----DPQT 114
 DB 121 ASGLFIQLIMLPTTICHAELERFLEHVDTSA-EDVAFPDIMWEVWKHKHGDPELL 179
 QY 115 LEQNTMDKNYMAHLYVEVQHERKASGQTFHLLTSLPRPRSTAPKPKSPQIQIG 174
 DB 180 SQFSMAHKYLPALDEPH-----PPKRLRSD--PDACPIMPLIAM 218
 QY 175 RIRVGTOL--RVLQPE-----DLAGMFIQFPIS-PDPR-----WQ 208
 DB 219 LLRGITQIOSRLTIGRGRCALANLADN-LYFALTEDDPQEVSTVYLDKATVTSWN 277
 QY 209 SSSPPRVALLQALQGLERARYVGGSPPEVGTIVRVIALATLLSPGALVMSHRSH 268
 DB 278 SDTONPYH--QALAE--KVKEARDVS-----LTSKLKPS-----ETI 313
 QY 269 FLACPLRLQLOQYQRCVODTGFSFLKVLLOMLQWLDSPGEGPLRAOLRMLASQAS 328
 DB 314 FVGCETLHLL-----LREW-----GEELOAVLRSSQGTSTY 343
 QY 329 AGRLSD-----VRGLRLAEALAFRODLEVS-----STV--RAV----- 363
 DB 344 DSYRLCDSLTSSQATLYLNTSLSKEDROYVSELAECVDFLKTSTYLNKRLIEDIT 403
 QY 364 -----IATLRSGQCSVEPDLISKVLOGLIE 389
 DB 404 ASIAMAIVIQKMDRHEVCYIFASEKWAFSDEWACAGSNALPREPDVLRLLLETVID 463
 QY 390 VRSPLHELLTAFSATADAAAPACKRVYVVSLLLOEPEPLAGGPGADGSGLEAVR 449
 DB 464 V-----STDRAIPESQIRQV--HLIECTADLS--LPGKN----- 496
 QY 450 LGPSSGLVMDLEMDPEVSSCPDLQLRLFSRRKGQ--AQVSPFRPYLLTF-- 503
 DB 497 -KVLGILRSM-----GKKSGSEKLAVVEGEQEDLNTTFNQL 533
 QY 504 THQSSWPTL-----HQCIRVLGSRDEQF-- 528
 DB 534 TQASAEGLAKAVASVARLYVHPETVKKMCSLAVNLTGHKFLQALITFPALRFEVY 593
 QY 529 -DPSASLDFLWACIHVPRIWOGRODQTPQKREELV-----RVQPELISLV-- 575
 DB 594 QCPNSSATFWSCLK-ETVWM--KFTREKEQFLELNLCLMSPVKPGQIYVAFLEPDE 650
 QY 576 --ELILAEATRSQDGTACSLIQ-----ARPLLLSCCGGDDSEVR 616
 DB 651 VLKEVLFPLRLDVEVDLSLRFIQTLEANNACREYMLQTCSPPLFLFSIC-- 702
 QY 617 KYTEHLSCIOQMGDSVYGRCRDILLQYLQRPRLRVPEVLLHSGAASSV-- 671
 DB 703 ---QLDRFSKYWPLPEKRC-----LSDRKDLAHIHELCEIYSANAFETSPDW 752
 QY 672 -----CKLDGLHFRITLLADYSDSLLENRGAD 700
 DB 753 IKSLSWLHRLKEQLDWTYGLRKSFEGHFKECVATPTEICKLSDEDETSOAHGYGAG 812
 QY 701 ASM-----ACRKLAIVAPLRLRLHLP--IAALLHGRTHLNFQEFROO 741
 DB 813 TGLLAMECCCVSSGISRMLSLVVDYGNPEVRLFSKGLFVALVYVWPMQSPQEMQRL 872
 QY 742 NHL-----CPHLVGLLELLQPHVFRSEHOGALMDCILSTRLLVYKSSRLAFL 795
 DB 873 HOLTRRLLEKOLLHVPYSLEYIQ-----FVPLP-----NLKFFAQL 909
 QY 796 NKEVOFIHKYITYNAPAIISFLQKHA--PLHDSFSDNSDVLTKSLAGLSLPSRSD 851
 DB 910 QLSVLFLELTF-----QPLCSHSCRNMPLBEKMH-----VKKLLCGSLT-- 948
 QY 852 RTDRGLDE-----EGEESASGSLPLVSVSLFTPLTAEMADYMKRLSRGQTV 900

Db 949 ---RLDVSRAIOAGFWGVPEDD-----LTQEALEFVYTOVFCALHTMAMLRPEVCE 999

QY 901 DL-----LEVLSIDIDEMSRRPETLSFFS-TNLORLMSAAEE 936

Db 1000 PLVVALFETITCYETLSKTPSVSSLQRAHQERFLKSTAE 1040

RESULT 15

VINC_CHICK

AC P12003: STANDARD: PRT: 1065 AA.

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Vinculin (Metavinculin).

GN VCL.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; OC Gallus.

OX NCBI_TaxID=9031;

RP SEQUENCE FROM N.A.

RC MEDLINE=89042216; PubMed=3141928;

RA Coutu M.D., Craig S.W.;

RT "CDNA-derived sequence of chicken embryo vinculin.";

RL Proc. Natl. Acad. Sci. U.S.A. 85:8539-8539(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA MEDLINE=89246350; PubMed=2497736;

RA Price G.J., Jones P., Davison M.D., Patel B., Bendori R., Gelger B., Critchley D.R.;

RT "Primary sequence and domain structure of chicken vinculin.";

RL Biochem. J. 259:453-461(1989).

RN [3]

RP SEQUENCE OF 1-880 FROM N.A.

RC TISSUE=Embryo;

RA MEDLINE=88024106; PubMed=3117046;

RA Price G.J., Jones P., Davison M.D., Patel B., Eperon I.C., Critchley D.R.;

RT "Isolation and characterization of a vinculin cDNA from chick-embryo fibroblasts.";

RL Biochem. J. 245:595-603(1987).

RN [4]

RP TALIN INTERACTION DOMAIN.

RA MEDLINE=90078327; PubMed=2512301;

RA Jones P., Jackson P., Price G.J., Patel B., Ohanion V., Lear A.L., Critchley D.R.;

RT "Identification of a talin binding site in the cytoskeletal protein vinculin.";

RL J. Cell Biol. 109:2917-2927(1989).

CC -1- FUNCTION: INVOLVED IN CELL ADHESION. MAY BE INVOLVED IN THE ATTACHMENT OF THE ACTIN-BASED MICROFILAMENTS TO THE PLASMA MEMBRANE.

CC -1- SUBUNIT: EXHIBITS SELF-ASSOCIATION PROPERTIES.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF ADHESION PLAQUES.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: VINCULIN (SHOWN HERE) AND METAVINCULIN. ARE PRODUCED BY ALTERNATIVE SPLICING. METAVINCULIN DIFFERS FROM VINCULIN BY THE INSERTION OF A 68 RESIDUES DOMAIN NEAR THE C-TERMINUS.

CC -1- PTM: PHOSPHORYLATED ON SERINES, THREONINES AND TYROSINES AND ACYLATED BY MYRISTIC ACID AND/OR PALMITIC ACID.

CC -1- SIMILARITY: STRONG, TO OTHER VINCULINS AND TO ALPHA-CATENINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.

DR EMBL: J04126; AAA49136.1; -

DR EMBL: Y00312; CAA68412.1; -

DR PIR: A27884; A27884.

DR PIR: A29997; A29997.

DR PIR: A31346; A31346.

DR PIR: S03973; S03973.

DR InterPro: IPR001043; Vinculin/catenin.

DR InterPro: IPR000633; Vinculin_2.

DR Pfam: PF01044; Vinculin; 1.

DR PRINTS: PRO0806; VINCULIN.

DR ProDom: PD025583; Vinculin_2; 3.

DR PROSITE: PS00663; VINCULIN_1; 1.

DR PROSITE: PS00664; VINCULIN_2; 3.

KW Cell adhesion; Actin-binding; Cytoskeleton; Structural protein; KM Phosphorylation; Repeat; Alternative splicing; Lipoprotein.

FT INIT MET 0

FT DOMAIN 167 207 TALIN-INTERACTION.

FT DOMAIN 258 588 3 X 112 AA TANDEN REPEATS.

FT REPEAT 258 368 1.

FT REPEAT 369 478 2.

FT REPEAT 479 588 3.

FT DOMAIN 836 877

FT MOD_RES 536 536 PRO-RICH.

FT MOD_RES 821 821 PHOSPHORYLATION (POTENTIAL).

FT CONFLICT 441 446 TAKLSD -> QLSQOI (IN REF. 2 AND 3).

FT CONFLICT 700 700 Q -> H (IN REF. 2 AND 3).

FT CONFLICT 879 879 E -> K (IN REF. 3).

SO SEQUENCE 1065 AA; 116867 MW; 8D6C887D9E71E77 CnC64;

Query Match 2.38; Score 121; DB 1; Length 1065;

Best Local Similarity 20.2%; Pred. No. 2.3; Mismatches 156; Gaps 26;

Matches 125; Conservative 101; Mismatches 236; Indels 156; Gaps 26;

QY 11 ILTLGPRRADDSFQALDIPWPEKKPLPTAFVDTSEALLDPLKLRMTRESERL 70

Db 121 LLTF-----DEAEVRKIIRVCKGILEYLVVAEYETMEDLVYTKNLGCGM--TKAKM 173

QY 71 VDAALQLEPQQ-----LLEVFSGIPVSSMSKLLQFLDOAVAHDPQL 115

Db 174 IDERQGLTQHEHRVMLVNSMNTYKELPLISMKLFVYTKNKSQGIIEALKNRHFTV 233

QY 116 EQNIMDKNYAHLYEV---OHERGASGGQTFHSLTASLPPRDSTE-----APKPS 165

Db 234 EKMSAEINIEIRVQLTSMEDMAMASKDTEAMKRALIDSKNKGAGWLRDPAAP-PGD 292

QY 166 SPEDPQCG-----GRIRVGTQLR-VLGPEDDLACMFQIPLSPDPKMOSSPPRPA 216

Db 293 AGEQAIQIIDEAGKAGELCAGKERREILGTCTLGQMTQDLADR--ARGQCATPMAMQ 350

QY 217 LALQALGOEL-ARVVOGSPVEVPIG-----VRVQALATLTLSPFGAL-----V 261

Db 351 KAQOVSGDLTLAKVNAARLKLEAMTNSKQATAKKIDAQNMLADPNNGSEGEHRTGI 410

QY 262 MSMSHRSHFLAC-----PLRLQLOCYQRCVPQDTGSSFLKVL---LQ 301

Db 411 MSEARKVAELCEEPKERRDLIRLSIGEISALTAKLSLRGRKGSPARALAQIATSLQ 470

QY 302 MLC-----WIDSPGVES-GPLRAQLMLASQASGR 332

Db 471 NLSKTRAVANTRPYKAAVHLEBKIEQAOQRWDNPVDRGVQAIRGLVAE---GRR 527

QY 333 LSDVRGGLRLALAEAFRODLEVYVSTVRVITLNS---GQCQSEPLILSLVLOGLI 388

Db 528 LANVMK-----PYRDLAKCKDRVQLAQLADLARAGESEQAATAINAQLODSL 579

QY 389 EVNSPHLELLTAFFS-ATADAASPPACKPVVAVSSLLQDEEPLA-----GKXP 438

Db 580 KDLKARQOEMTQEVSVDFSDTTPIKLAVVATAVSDPVRNREVEEERAAENFENHARL 639

QY 439 GAQGSQLEAVRLRPPSSSLVLDMLDEM-----LDPEVSSCPDLQLRLFSRRKKGQAO 491

Db 640 GATAKAAAV--GTANKTVEGICATVKSAREITPOVSA-----RIL--RNPENQAA 690
QY 492 VPSFRPYLLTFTHOSW 509
| |
Db 691 YEHE-----TKNQW 701

Search completed: May 7, 2003, 19:41:31
Job time : 38 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:40:03 ; Search time: 19 Seconds
(without alignments)
1593.484 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 5241
Sequence: 1 MHILVHAMVILLTGPRA.....MDPSAQISEALRLHMEAVM 1029

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	136.5	2.6	2475	US-09-413-814-48	Sequence 48, Appl
2	133	2.5	737	US-08-188-582-16	Sequence 16, Appl
3	133	2.5	737	US-08-646-715-16	Sequence 16, Appl
4	126	2.4	878	US-08-708-541A-26	Sequence 26, Appl
5	124.5	2.4	5087	US-09-144-085-1	Sequence 1, Appl
6	124	2.4	577	US-09-413-814-4	Sequence 4, Appl
7	123.5	2.4	1455	US-08-726-012B-2	Sequence 2, Appl
8	121	2.3	1315	US-09-031-563-2	Sequence 2, Appl
9	121	2.3	1315	US-09-031-563-25	Sequence 25, Appl
10	121	2.3	1315	US-09-293-505-10	Sequence 10, Appl
11	121	2.3	1315	US-09-392-277-2	Sequence 2, Appl
12	121	2.3	1315	US-09-392-277-25	Sequence 25, Appl
13	120	2.3	3066	US-08-952-127-12	Sequence 12, Appl
14	119.5	2.3	819	US-09-651-656-15	Sequence 15, Appl
15	119.5	2.3	819	US-09-650-855-15	Sequence 15, Appl
16	116.5	2.2	3798	US-09-335-409-6	Sequence 6, Appl
17	116.5	2.2	3798	US-09-568-102-6	Sequence 6, Appl
18	116.5	2.2	3798	US-09-567-969-6	Sequence 6, Appl
19	116.5	2.2	3798	US-09-568-480-6	Sequence 6, Appl
20	116.5	2.2	3798	US-09-568-486-6	Sequence 6, Appl
21	116.5	2.2	3798	US-09-568-472-6	Sequence 6, Appl
22	116.5	2.2	3798	US-09-567-899-6	Sequence 6, Appl
23	115.5	2.2	1055	US-09-031-563-27	Sequence 27, Appl
24	115.5	2.2	1055	US-09-392-277-27	Sequence 27, Appl
25	115	2.2	1012	US-08-811-481-16	Sequence 16, Appl
26	115	2.2	1621	US-08-242-677-2	Sequence 2, Appl
27	114	2.2	979	US-08-514-213A-2	Sequence 2, Appl

28	114	2.2	1208	US-09-463-702A-2	Sequence 2, Appl
29	112.5	2.1	804	US-08-909-954-2	Sequence 2, Appl
30	112	2.1	3072	US-09-413-814-93	Sequence 93, Appl
31	112	2.1	3079	US-09-413-814-80	Sequence 80, Appl
32	111	2.1	1004	US-08-916-352-2	Sequence 2, Appl
33	111	2.1	2627	US-08-751-189-3	Sequence 3, Appl
34	111	2.1	2627	US-09-060-836-3	Sequence 3, Appl
35	111	2.1	2627	US-09-184-445-3	Sequence 3, Appl
36	111	2.1	3056	US-08-508-836A-8	Sequence 8, Appl
37	111	2.1	3056	US-08-629-001A-3	Sequence 3, Appl
38	111	2.1	3056	US-08-874-266-2	Sequence 2, Appl
39	111	2.1	3056	US-08-642-274D-3	Sequence 3, Appl
40	111	2.1	3056	US-08-952-127-3	Sequence 3, Appl
41	111	2.1	3056	US-08-952-014C-3	Sequence 3, Appl
42	111	2.1	3056	US-09-360-416-2	Sequence 2, Appl
43	110.5	2.1	3057	US-09-360-416-3	Sequence 3, Appl
44	109	2.1	1745	US-09-031-485-33	Sequence 33, Appl
45	109	2.1	1745	US-08-847-429A-33	Sequence 33, Appl

ALIGNMENTS

```
RESULT 1
US-09-413-814-48
; Sequence 48, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Beyer, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin Ver. 2.1
; LENGTH: 2475
; TYPE: PR
; ORGANISM: Sorangium cellulosum
US-09-413-814-48

Query Match
Best Local Similarity 20.7%; Score 136.5; DB 4; Length 2475;
Matches 193; Conservative 124; Mismatches 341; Indels 275; Gaps 50;

QY 178 VGTQLRVLPEDDL-----AGMFLQTFPLSPDRMSSRPVALALQALGCELAR 229
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 785 IGDPLPAGAGAEEDLLAASASAGS-----PPESPANSAMERPRAOS-----SIAS 830
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 VVQGSPEVGIYRVQLATLISPH---GGALVMSHRSHEFLACPLRLQCOYRCVP 286
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 831 ATPPASHSASV---AAATILEYRAYFGFAAVSTD-AFF----- 868
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 287 OPTGESLFLKVLQMLQMLDSFGVE-----GGFLRAQLRLMASQASAGRRLSDVRGG 339
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 869 -ELGASSLDLVNIGQLS--DRIGREVPTLLLYDHTPPQALALTSALSAEAPPLRGG 925
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 340 LRLALALAFRODLEVYSTRAVATLTLSRGEQCSVEPPLISKVGLIEV--RSFHLER 397
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926 -----HRASTSGTASASASTAPTPGDAHS--QPSFVREODIAITGMFRGCGADD 975
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 398 LTAFTSATDAASPPACKPVVSSLLQEEBPLAGKPGADGSGLEAVRLGSSGL 457
D 976 -LDAFNNLVEG-----VESITPFSEDELLAGVPREHASTRYVR--ANGEL 1020
QY 458 VDMLEMDPEVSSCPDQLRLLFSSRRKKGQAVPSFRYLLTLFFHOSWPTLHCIR 517
D 1021 TGAMD-PEPE-----FFGTSARAAVMDPOFRVF-----HECSMAL----- 1056
QY 518 VLLGKSREORFP--SASLD-FLMACIHPRLMOGRDQRTPOKREELVLR----- 565
D 1057 -----EHGYDDPTPCASIGVAGVYNNHLPWLMRLPLHLEBQFALLTLTREFFAPL 1110
QY 566 -----VQGEEL-----ISLVELLAEAFETSQSDTACSLIOARLPLLSCCG-- 610
D 1111 LSYKVGIRGPAISLOFACSTSLVAIGTACBELRA-----GACQMALAGVTASIERGCFE 1165
QY 611 -----DDESVRYKTEHLSGCGQMGDSVLGRRCRDLLQLYLRPELVRPPEVLH 662
D 1166 HOGGYILSPGHTRSFDPAAAGTV--FGDGV-----GMVLK-----PLAQL-- 1206
QY 663 SEGASSSVCKLDGLHRTFTLLADTSDSBALENRG-----ADASMKRLAVAHPLLL 717
D 1207 ADDDTIAVAKGIGINNDAKRGVFTAPSRAGOTEATRALRDAGVANSRYVE----- 1261
QY 718 RHLPMIALHGRTHLNFQEFROONHLSCLHVLGLLELLQPHVRESEHOGAL--WDCLL 775
D 1262 -----AHGTA-----TRMGDPIVEALTO-----AFRAEADGPRLPGSCLL 1297
QY 776 SFIRLLNYSKSSRHIAFINKFQVFIHKYIT--YNAP-AAISFQKADPLHDSFD 830
D 1298 GSKSVNGVHNAAGVAGLVKTVLALOHRLPISLFYQSNPHDFA--ASPR-ANGQ 1353
QY 831 NSDLVMLKS--LLAGLSLPSRDRDTRDGLDEGESESSAGSLP----- 872
D 1354 TSDMAVEGRRLAGVSSFIGTNAHLIYEAPKALPTTAALSTEPNDLADGADGLV 1413
QY 873 VVSLEFPLTAEMAPRM-KRLRGQVEDELLEVLSDID--EKSRR-PELLSFESTNL 927
D 1414 LPIASAPPLAHIAHIANLHNLHPRI-----ALADVALTLQGRQWPHRSLI---- 1464
QY 928 QRLMSAECCRNIAFSL-ALRSMONSPSIAAFLP-----TFWV-CGSGDFEVVOT 978
D 1465 -----CRNRTEIKILRAVHSAEVPRAQAPVSDAPRCVFLFPGGAGVPSMARD 1514
QY 979 ALRNLPEVAL--LQGEHAVAL--HRAFLVG 1005
D 1515 LVNCPDFALHLDPCLDQLAELLPEDPICILFG 1547

```

```

RESULT 2
US-08-188-582-16
Sequence 16, Application US/08188582
Patent No. 5534410
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLHR, HOHACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435.
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ. ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 737 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-188-582-16

```

```

Query Match 2.5%; Score 133; DB 1; Length 737;
Best Local Similarity 20.8%; Pred. No. 0.00055;
Matches 135; Conservative 78; Mismatches 233; Indels 202; Gaps 28;

```

```

QY 135 RGASGGTFHSL-----LTASLPPRRDSEAPKPKSSP-----QDPIGGGIRV 178
D 54 KGAAVAT--OSSLRTPTATTSIGIRATLTPVLAERLPQNPNTIONFOLPREMWLVRS 112
QY 179 GTQLRVLPEDDLAGMFLQIF-----PLSPDPNROSSPPRVALALQALGOE-LARVQ 232
D 113 ENGLLMIPQALALQMAQOAHQAPQTMARPRATPTSAP--PVGISVQAPGTPIANQVT 171
QY 233 GSPVPGIT--VRYLALATLLSPH-----GVALVSMHSHFLACLLRQLQCOYR 283
D 172 PTLIKVQSOAQTTVPATLQSPGVOPQLVYGAA-----QTASLGTAVQGTGPR 226
QY 284 CVPDDTFESSL-----FLKVLQMLQWLDSPVEGGPRLAOLRMLASQASAG 330
D 227 TVGATITSSAATETMENVKCKNFSLTIK--LASSKOSTETVANKVELYONLIDG 282
QY 331 R-RLSDVRGGLR-----LAEALAFRDLE----- 354
D 283 KIEAEDFTSLYRELNSPPQVLPFLKRLPALRQLTPDSAAFIOQSOQPPPTSOAT 342
QY 355 -----VVSSTVR-----AVATILRSGEQCSY-----EPDLISKVLQG----- 386
D 343 TALTAVALSSVORTAGKTATATYSALQPPVLSLQPTQVGVGQGPTELVIOQPPKG 402
QY 387 -----LIEVRSPHLEELITAFSAATADASPPACKPVVSSLLQDEEP 432
D 403 ALIRPQVTLTQTPMAVLRQPHNRIMLTTPQOQLNPLOVPVYKRAVLDEGTALSAVSA 462
QY 433 LAGG-----KPGA-----DGSLEAVRLGSSGLVDWLEMDPEVSSCPD- 474
D 463 QAAAGKRLKEPGGGSEFRDDDDINDVASWAGVNLSESRILATINSELVGTLTRSKDE 522
QY 475 -----LQRL-L-FSRKRGKGAQVPSFRYLLTLFTHOSWPTLHCIRVLLGK---- 522
D 523 TFLQAPLQRLRIIEIGKHH-----ITELHPDVAVSYSHAT-----QORLQNLVEKISFT 572
QY 523 -----SREORFDSASLDFLMACIHPRLMOGRDQRTPOKREELV-----L 584
D 573 AQQKNSYKDDRYEQASDVRAQLKFF-----EQDLQIKQKKDEBERIILIRAKSRS 626
QY 565 RVQGEPLISL-----VELLIAE-AETRSQDGTACSLIOARLPLLSGC 607
D 627 ROEDPEQLRLKQAKAKEMQOELAQMRQDANLALAAIGRRKKRYKVC 674

```



```

RESULT 3
US-08-646-715-16
: Sequence 16, Application US/08646715
: Patent No. 5637686
: GENERAL INFORMATION:
: APPLICANT: Tjian, Robert
: APPLICANT: Comai, Lucio
: APPLICANT: Dynlacht, Brian D.
: APPLICANT: Hoey, Timothy
: APPLICANT: Ruppert, Siegfried
: APPLICANT: Tanese, Naoko
: APPLICANT: Wang, Edith
: APPLICANT: Weinzierl, Robert O.J.
: TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
: TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/646,715
: FILING DATE: 09-MAY-1996
: CLASSIFICATION: 435
: PRIORITY INFORMATION:
: APPLICATION NUMBER: US 08/188,582
: FILING DATE: 28-JAN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 737 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-646-715-16

Query Match 2.5%; Score 133; DB 1; Length 737;
Best Local Similarity 20.8%; Pred. No. 0.00055;
Matches 135; Conservative 78; Mismatches 233; Indels 202; Gaps 28;

QY 135 RGASGGGTFHSL-----LTFSLPFRNRSTAPPKKSP-----EOPIGGRIRV 178
Db 54 KGAACAVT-QSLSRPTATTSIRATLTPYLAAPLPDPONPTIIONFQLPBGAVLRS 112
QY 179 GTQLRVLPEDDLACMLQIF-----PLSPPRMOWSSPRVALAQQALGQE-LARVVO 232
Db 113 ENGQALMIPOQALQMAQNAHQNPOTMAPRAPPTTSAP-PIQISTVQAPGPIIARQVT 171
QY 223 GSPEVPGIT--VRVLOALATLISPH-----GQALYMSHRSFPLACPLRLQCLQYOR 283
Db 172 PTTIKQVSOAQTVQPSATLQRSRGVOPQVLVGGAA-----QTASLGTATAVQTGTQOR 226
QY 284 CVPQDTGFSLSL-----FLKVLQMLQMLDSPGVCGLRLQRLMLASQASAG 330
Db 227 TVPGATTSSAATETMENVKCKNPLSTLIK-----LASSGKOSTTAANVKELVONLIDG 282
QY 331 R-RLSDVKGGLLR-----LAELAFQDLE----- 354

```

```

Db 283 KIEADFTSRLYRELNSSPQYLPVLPFLKRSLPALROLFPDSAAFIQSSQOQPPPTSOAT 342
QY 355 -----VVSIVR-----AVIATLRSGEQCSV-----EPDLISLVNG----- 386
Db 343 TALTAVLVSSSVQRTAGTKATATVTALQPPVLSLTQPPVOVGKQGPPTPLVQPPKRG 402
QY 387 -----LIEVSRPHLEELTAFFATDASPPACKPVVYVSSLLQEEEP 432
Db 403 ALIRPPQVTLTQTPWALRQPHNRIMLTTPQOIQLPQVPVYVPAVLPKALSAVSA 462
QY 433 IAGG-----KPGA-----DGSLAVRLGPPSSGLVDLWLEMDPEVSSCPD- 474
Db 463 QAAAQKNNKKEPGGSPRDDDDINDVASMAGVNSEBARILATNSELVGLTRSKDE 522
QY 475 -----IQRLIL-FSRKRGKQAOVPSRPILLFLFTHQSSWPLHQCIRVLGK----- 522
Db 523 TFLQAPLQRLILEIGKRG-----ITELHPDVSVSHAT-----QORLQNLVERKISPT 572
QY 523 -----SREQRPDSASLDPLMACIHVPRIWGRDQRTPOKRREELV-----L 564
Db 573 AQQKNFSYKDDRYEQASVVRQLKF-----EQLDQLEKRGKDEKREILMRKAKRS 626
QY 565 RVQGPFLISL---VELILAE-AETRSODGDTAACSLLQARLPPLLSC 607
Db 627 RQDEPEQLRKQKAKEMQOQELAQMRDANLTALAIQPRKKRYKDC 674

RESULT 4
US-08-708-541A-26
: Sequence 26, Application US/08708541A
: Patent No. 5871744
: GENERAL INFORMATION:
: APPLICANT: VAKHARIA, Vikram N.
: APPLICANT: MUNDT, Egbert
: TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIKAIKO, MARKELSTEIN, MURRAY & ORAM LLP
: STREET: 655 Fifteenth Street, N. W.,
: CITY: Suite 330 - G Street Lobby
: STATE: DC
: COUNTRY: USA
: ZIP: 20005-5701
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/708,541A
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: KITTS, Monica C.
: REGISTRATION NUMBER: 36,105
: REFERENCE/DOCKET NUMBER: P8172-6002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202/638-4810
: TELEFAX: 202/638-4810
: INFORMATION FOR SEQ ID NO: 26:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 878 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-708-541A-26

Query Match 2.4%; Score 126; DB 2; Length 878;
Best Local Similarity 19.9%; Pred. No. 0.0038;
Matches 155; Conservative 104; Mismatches 279; Indels 242; Gaps 39;

```

[illegible][illegible]

Db 382 QEVHMOVLSFVSALVCPEDAOQLLEDMVWRLMAQAFESCCQDLSMTAVLVRQAALG 441
 QY 112 PQTLEQNMKMYMAHLEVOHERGASGO-----TFHSLTASLPERRDSTE 159
 Db 442 PSA-----FLSYADMWRKASGSTRGHGCKKALVLEFTLSLVEPESRYLOVH 492
 QY 160 APPKSSPEQ-----PIQGRIRVG---TOLRVLGPEDDLGMFLQIPELSPDRMOS 209
 Db 493 ILHPPLVPSKRSRLTDYISLAKTRLADLVKYSIENMGLYEL----- 534
 QY 210 SSPRVVALQOALGOELAVYOGSPVPGITVRVLOALATLLSSPHGALVMSHRSHF 269
 Db 535 SSAGDIEPHSOAL-QDVEKAIMVEHTGNIPYVME--ASIFRRPY-----YVSHF 583
 QY 270 LACPRLRLQOYRCYVPODTGFSSLFKVL-----IQMLQMLDPSGVGGPPLRAOL 320
 Db 584 LPALLT-----PVLKPKVDSRVAFIESLRADKIPPSLYSTVCOQSAEKPEDAL 637
 QY 321 RMLASQASAGRRLSDVAGGLRLAELAFRODEVSVSTYRAVATYTR-----SGECCSV 375
 Db 638 GVRAPENSAEPLGOLTAALGELMASMTDPSQRDVISAQVAVISERLRAVLGHNEDDSV 697
 QY 376 EPDLISKVLGLIEVRSPHLEE-----LTAFFSATADAASPP----- 414
 Db 698 E-----ISKI---QLSINPRLEPREHIAVDLLTSFCOMLMAASVAPPEKOGPAALFVR 751
 QY 415 -ACPRV--VVSLL--LLOEEPLAGKPGADGSLAVRLGPSSGLLVMLEMLDEVV 469
 Db 752 TMCGRVPAVLTRLCQILRRHQGPLSA-PHYGLAALAVHIGESRSAL-----BEVD 802
 QY 470 SSCPDQLRLIFSRKKGGAQVPSFRPYLLTLFTHOSSMPTLHOC---IRVLLGKREQ 526
 Db 803 VGPR-----APAG--LVPALFDSLTCRTDLSFCLKFCYAAISYSCLKRESSQ 851
 QY 527 RFPSPAS-----LDPL-----MACIHP-----RIV 547
 Db 852 SRDILCSLPGLIKKQFLMFRLFSAROPLSSEDAVSLSMRPLHPSADWQMAALSIW 911
 QY 548 QGRDQTPQKRRE-----ELVRYQGPBELISLV-----ELIAEAETR 585
 Db 912 THRFREVLKEEDVHLYTQDMLHLELIO--PEADALSDTERODEHQAIAHHEFLPRESSA 970
 QY 586 SQ--DGDVACSLIOARPLLLSCCCGDESVRKYTEHLSCIQMGSVLGC----- 635
 Db 971 GCGDGD-----LQACTITVNALMDFHOSRST-DH-----SNSDLVFCGRTGNEDI 1018
 QY 636 RCRDILLQYLRPELRLVPEVVLHSEGAASSVCKLDGLIHRFTLLADTSDSRAL 695
 Db 1019 SRLQEMVADLEQO-DLIVPL-----GHTPSQEHFLFEIFRRRIQALTL-----S 1061
 QY 696 NRGADASMACKLAVAHPLLLRLHPLMIALHGRTHLNFOEFRQONHSC---FLHVL 751
 Db 1062 GWSVAASLQORELIMYKRIILR-LP--SSVLCG-----SSFOAEOPITARCEQFPHLV 1112
 QY 752 GLELLQPHVFRSE-----HOGAL-WDCLLSFRLILNRYKRSRHIAAFINKVOGFIHK 804
 Db 1113 -----NSEMRFCSHGALITODITAHFFRGLNACLKSRDPSL---WVDILA 1157
 QY 805 YITYNAPAAISFLQKHAADPLHLSFDNSDLVMLKSLAGL-----SLPSRDRTD 854
 Db 1158 KQCTKPLILT-----SALVWMPSLPEVILLCRWRRCQSLPRELQKL- 1200
 QY 855 RGLDEGESESSAGSLPLVSVSLFTPLTAEMAPYMKRLSRGO---TVEDLLE--VLSDI 909
 Db 1201 -----QOEGRQFAP-----FLSPEASAPAPNPMDLSAAALHFAIQOVRREENIKROL 1246
 QY 910 DEMSRRPPELLSF-FSTNLOQLMSSAECCRNLAFLSALSMONSPS---IAAFLEPTF- 964
 Db 1247 KKLDCEREELLVLEFFLSLGLLSS-----HLTSNSTDDLPRAFHYCAAILCCL 1296
 QY 965 -----MYCLSGQDEFEVQFALRNLP-----YALLCQOEHA-VLLCHRAFL- 1003
 Db 1297 KRIISMLALPQLTESDRLRLRLRVAPDOHTRLLPAPYSLLSYFHEDAIIREEAFLHV 1356

QY 1004 -VGMY 1007
 Db 1357 AVDMY 1361
 RESULT 8
 US-09-031-563-2
 ; Sequence 2, Application US/09031563A
 ; Patent No. 6022708
 ; GENERAL INFORMATION:
 ; APPLICANT: Frederic de Sauvage
 ; APPLICANT: Arnon Roshenthal
 ; TITLE OF INVENTION: Fused
 ; FILE REFERENCE: P1272
 ; CURRENT APPLICATION NUMBER: US/09/031,563A
 ; CURRENT FILING DATE: 1998-02-26
 ; NUMBER OF SEQ ID NOS: 27
 ; SEQ ID NO 2
 ; LENGTH: 1315
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-031-563-2
 Query Match 2.38: Score 121, DB 3: Length 1315;
 best Local Similarity 20.08: Pred. No. 0.025;
 Matches 243; Conservative 130; Mismatches 426; Indels 414; Gaps 54;
 QY 44 LVOTSEBALPLDMLKRMIRSEVRLVDAALDLEPOQLLTFVQSGIPVSS-MSKILQ 102
 Db 235 LTRKPRRLSPDLLYHPFLTAGHVTITTEPAGPDL-----GTPFTSRLLPELO 282
 QY 103 FLDOAVAH--DPTLEQNMKMYMAHLEVOHERGASGQTFH----- 144
 Db 283 VLKDEQAHLRLAPKNGQSRILITQAYKRMAEAMQKKNHTGPALEQEDKTSKAPGTAPLP 342
 QY 145 -----SLL-----TASLR--PRDSTEAPKPSPEQPIQGRIR 177
 Db 343 RLGATPOESSLILAGILASELKSSWAKSGTGEVPSAPRENRTTPDCERAPPE----- 394
 QY 178 VGTOLRVLPED-DLACMFQIPLSPDPWQ-----SSSPRVALL----- 217
 Db 395 ---RPEVLQGRSTVDV--LENEPDSDNEMQHILLETTEPPIQKAPRLTILCPDCOR 449
 QY 218 ---ALQOALGOELARVVOGSEVPVITVRVLOALATLLSSPHGALVMSHRSHFLACPL 274
 Db 450 IQSOLHEAGGQILKGILEGASHI---LPAFRVLSLSSCSDSVALYSECREAGLPLGL 505
 QY 275 LRQL--GOYQRCVPODTGFSSLFKVLQMLQMLDPSGVGEGPLRAQLRMLASQAS---- 328
 Db 506 LSLRHSQESNSLQOQSMYGT-FLQDLMVAIYQ-----AFACFTFNLERSQTSLSQ 555
 QY 329 ---AGRRLSDVRGSL-----LRLAELAFRODLEVSVSTYRAV---LATLRSG 370
 Db 556 VFOGAANLFDLDLGLKLAQPDSDSQTLRBSIMCFYVLCAMODNSAISAAPYSSLLTT 615
 QY 371 EOGSEV-----PDLISKVLGLIEVRSPHLEE-----LTAFFSATADAASPPAC 416
 Db 616 QQVVLDELHGLTVPQLPVHTPGAPQVSOPLRQSESDIGALISSALAICTAPAVGLPDC 675
 QY 417 -----KPY-----VYVSSLLQGE-----EE 431
 Db 676 WDAKEQVCWMLANOLTEDSSQLRPSLSIGLQHPILCHLKLKVLVSCCLVSGICRLGLQ 735
 QY 432 PLAGGKRGADGSGLEAV-RLGPPSSGLLVDMLEMDPEVSSCPDLQRLRLFSSRKKG--G 488
 Db 736 PLA-----LESPLMLQGVKVVYDWESEVTLTYF-----LSLVFRLQNLPCG 779
 QY 489 QAQVPSFRPYLLTLFTHOSSMPTLHOCIRVLLGKSRQ----- 526
 Db 780 MERKGS--DVATLFTH-SHVSVLSVAACUQLGLQGGVTFDLQPMWMAAATHALSAP 835
 QY 527 ---RFPDSASLDPLMACI-----HVPRIWQGRDQTPQKRREBELVLRVQGP 570

```

Db 836 AEVRLTPGSCGFYDGLILLQLLLEQKASLIRDMSSSEKMTVLMHRSKVLRL--PE 893
Qy 571 LISLV--ELLIAEAETRSOD-----GDTAACSILQARL-----PILLSCCGGDESEVRK 617
Db 894 EASAOEGELSSPPSPPEPMTLISPPGMAALLSLAMATFTGEPOLCISCSOHGSIIMS 953
Qy 618 VTEHLSGCIQOWGDSVLGRCDLLQLYLQRPRLRVPEVYLHSEGAASSVCK---- 673
Db 954 ILKHL-----LCPSPFNQL--RQAPHSSEFLPVVYL-----SVQQLCF 990
Qy 674 ---LDGLIHRFTLLADTSDSRALENRGADASMACRLAVAPHL--LLRHLPMIAALLHG 729
Db 991 PFALDMDADLLIVLADLRDSE-----VAHLIQVCCYHLEPLMQVELP- 1033
Qy 730 RTHLNFQERQONHLSCLFHVIGLELLQPHVFRSEHOGALMDCLSFIRLLLNKRSR 789
Db 1034 -----ISLITRLALMDP-----TSLNPFVNTVASPR 1060
Qy 790 HLAEIFNKVQFIHKYITYNAPPAISFLQKHADPLHDSFDSNDLVMLKSLAGLSLPSR 849
Db 1061 TIVSFLSVALSDOPLTSDLSLTAHTARVLSFSH--LSF-----IQELLAGSDESTR 1112
Qy 850 DDRTRDGLDEGEESSESSAGSLPLVSLSFTPLTAAEMAPYMKRLSRGQTVEDLLEVLSDI 909
Db 1113 PLRSLIGHPE-----NSVRAHTYRLIGHLLQHSMLRGALQOSGILLSLLGLGDK 1164
Qy 910 DEMSRRREILSFSTNQLRLMSSAECCRNIAF--SLALRSQNSPSTAAFLPFMVC 968
Db 1165 DPVAV-----CSASFAVGNAAYQAQPLGAPALAA--VPSMTOLL 1201
Qy 969 GSODEFV---VOTALRN-----LPEYALLCOEHAVALLHRAFLVGMYGOM--DPSAQIS 1017
Db 1202 GDPQAGIRRNVASALGNIGPELGELLQCEVPRQL-----EMACGDQOPNVK 1250
Qy 1018 E---ALRIHME 1026
Db 1251 EALIALRSLQOE 1263

RESULT 9
US-09-031-563-25
; Sequence 25, Application US/09031563A
; Patent No. 6022708
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage
; APPLICANT: Arnon Koshenthai
; TITLE OF INVENTION: Fused
; FILE REFERENCE: P1272
; CURRENT APPLICATION NUMBER: US/09/031.563A
; CURRENT FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 25
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence
US-09-031-563-25

Query Match 2.3%, Score 121; DB 3; Length 1315;
Best Local Similarity 20.0%; Pred. No. 0.025;
Matches 243; Conservative 130; Mismatches 426; Indels 414; Gaps 54;

```

```

Db 343 RLGATPQSSSLLAGLILASELSKMSAKSGTGEVSAAPREKRTPDCEARPEE----- 394
Qy 178 VGTQRLVLCPEP-DLAGEFLQIFLSPDRMO-----SSSPRYAL----- 217
Db 395 ---REVILQORSTDAVD--LENEEPDSNEMOHLLETTEPEVITQKAPRTLLCNPDFCOR 449
Qy 218 ---ALQOALGOELAVNOGSPPEVPGITVAVLOALATLLSSPPGALVYMSHRSFLACL 274
Db 450 IOSLHEAGQILKILBEGASHI-----LPAFVYSSLSGSDSVALSFCREAGLPGL 505
Qy 275 LKQL--COYORCVPODTGESSLFLKVLQMLQMLDPSVEGSGPLRAOLMMLASQAS--- 328
Db 506 LSLRHSQESNLSQQQSWGT--FLQDLMAVIO-----AFVFACTFLERSQJSDSLQ 555
Qy 329 ---AGRLSDVRGGL-----LRLAEALARQDLEVVSVTRAV---IATLNSG 370
Db 556 VQEAANLFLDLGKILQPDDEBQTLRRDSIMCFVLCENADGNSRAISKAFYSSLLTT 615
Qy 371 EGCYSVE-----PDLISKVQLGLEVRPHLE-----LTLAFSATADAAPPPAC 416
Db 616 QOVVLDGLLGLTVPOLPVHTPOGAPQVSOPLREQSEDIPGAISALAICTAPVGLPDC 675
Qy 417 -----KPV-----VVVSSLLQOE-----EE 431
Db 676 WDAKEQVCWHLANQLTEDSSQLRPSLISGLQPHIICLHLKLYSCCLYSEGLCRILQGE 735
Qy 432 PLAGKPGADGSLAV--RLGSSGLYDWMLEMDREYVSSCPDQLRLFRKRGK--G 488
Db 736 PLA-----LESLFMLIOGKVVVDMESTETLYF-----LSLIVERLQNLPG 779
Qy 489 QAQVSPFRPILFTLTHQSWPTLHQCIRVILGKREQ----- 526
Db 780 MKRLS---DVATLFTH--SHVSYLSAACLQGLQCGQCVTTDIDPMEMAATIALSNP 835
Qy 527 ---RPDSASLDFLMACI-----HVPRIWGRDQRTPOKREELVLRVQGE 570
Db 836 AEVRLTPGSCGFYDGLILLQLLLEQKASLIRDMSSSEKMTVLMHRSKVLRL--PE 893
Qy 571 LISLV--ELLIAEAETRSOD-----GDTAACSILQARL-----PILLSCCGGDESEVRK 617
Db 894 EASAOEGELSSPPSPPEPMTLISPPGMAALLSLAMATFTGEPOLCISCSOHGSIIMS 953
Qy 618 VTEHLSGCIQOWGDSVLGRCDLLQLYLQRPRLRVPEVYLHSEGAASSVCK---- 673
Db 954 ILKHL-----LCPSPFNQL--RQAPHSSEFLPVVYL-----SVQQLCF 990
Qy 674 ---LDGLIHRFTLLADTSDSRALENRGADASMACRLAVAPHL--LLRHLPMIAALLHG 729
Db 991 PFALDMDADLLIVLADLRDSE-----VAHLIQVCCYHLEPLMQVELP- 1033
Qy 730 RTHLNFQERQONHLSCLFHVIGLELLQPHVFRSEHOGALMDCLSFIRLLLNKRSR 789
Db 1034 -----ISLITRLALMDP-----TSLNPFVNTVASPR 1060
Qy 790 HLAEIFNKVQFIHKYITYNAPPAISFLQKHADPLHDSFDSNDLVMLKSLAGLSLPSR 849
Db 1061 TIVSFLSVALSDOPLTSDLSLTAHTARVLSFSH--LSF-----IQELLAGSDESTR 1112
Qy 850 DDRTRDGLDEGEESSESSAGSLPLVSLSFTPLTAAEMAPYMKRLSRGQTVEDLLEVLSDI 909
Db 1113 PLRSLIGHPE-----NSVRAHTYRLIGHLLQHSMLRGALQOSGILLSLLGLGDK 1164
Qy 910 DEMSRRREILSFSTNQLRLMSSAECCRNIAF--SLALRSQNSPSTAAFLPFMVC 968
Db 1165 DPVAV-----CSASFAVGNAAYQAQPLGAPALAA--VPSMTOLL 1201
Qy 969 GSODEFV---VOTALRN-----LPEYALLCOEHAVALLHRAFLVGMYGOM--DPSAQIS 1017
Db 1202 GDPQAGIRRNVASALGNIGPELGELLQCEVPRQL-----EMACGDQOPNVK 1250
Qy 1018 E---ALRIHME 1026

```


QY 145 -----SL-----TASLP--PRDSTEAPKPKSSPEOPIGGRIR 177
Db 343 RLGATPOESSLAGILASSELKSSMAKSGTEGVSAPRENTTTPDCERAFPEE----- 394
QY 178 VGTOLRVIGED-DIAGMFQIIFPLSPDRMO-----SSSRPAL----- 217
Db 395 ---RPEVLGGRSDVD--LENEPDSDNEMOHLLETTEVPIQLKAPLTLICNPFCOR 449
QY 218 ---ALQALGOELARVVGSPPEVPGITVRVLOALATLSSPHGALVMSHRSHFLACPL 274
Db 450 IQSOLHAGGQILKILLEGASHI-----LPAFVLSLSSCSDSVALYSCREAGLPGLL 505
QY 275 LRQL--COYRCVPQDTEGSSLEFLKVLLOMLDMSPGVGGPLRAOLRLAQAAS----- 328
Db 506 LSLRHSQESNSLQOOSWYGT-FLQDLMAYIQ-----AVFACTFNLERQTSDSIQ 555
QY 329 ---AGRLSDVVRGL-----LRLAELAFRODDEVVSSTYRAV-----IATLRSG 370
Db 556 VFOEAMNLFIDLGLKLLAQPDSEQTLRRDSLKCFVLCAMGNSAISKAFYSSLLTT 615
QY 371 EOCSEV-----PDLISKVLOGLEVRSPHLE-----LITAFSATADAASPPAC 416
Db 616 QQVVLGDLHGLTVPOLPVHTPOGAPQVSPLEQSEDIGALISSALAICTAPVGLPDC 675
QY 417 -----KPY-----VVVSSLLQOE-----EE 431
Db 676 WDAKEQVCWMLANQLTEDSSQLRPSLISGLQHPIILCHLKLKLVYSCCLVEGCLRLQOE 735
QY 432 PLAGKPGADGSLLEAV-RIGPSSGILLVDMLEMDPEVSSCCDLOLRLEFSRRKG--G 488
Db 736 PLA-----LESIFMLIQCKVKVYDWEESTVTLVF-----LSLLVFLQNLPCG 779
QY 489 QAOVPSFRPYLLTLFTHOSSWPTLHOCIRVLLGKSRQO----- 526
Db 780 MERIGS--DVATLFTN-SHIVSVLSVAACLLQLGQOQVTFDLOPMEMMAATHTALSAP 835
QY 527 ---RFPDSASLDFLMACI-----HVPRTMOGDKQTPROKREBELVLRQOPE 570
Db 836 AEVRLTPPGSCGYDGLILLQLLQTEOGKASLIRDMSSSEMVTYLMHRSNMLRT--PE 893
QY 571 LISLV--ELILAEATERSOD-----GDFAACSLIQARL-----PILLSGCCDDSEVR 617
Db 894 EASNGEGLSLSPSPSEPPMTLISPOGMALLSLAMATTOEPOLCLOSLGSHGLMS 933
QY 618 VTEHLSGCIQOWGDSVLRRCRDLLQLYLQREPLRVPRVEVLHSEGAASSVCK---- 673
Db 954 ILKHL-----LCPSFLNQL-RQAPHGSEFLRVVL-----SYCQLLCE 990
QY 674 ---LDGLIRFTLLADTSSSRLBNRGADASMAKRLAVALPL-LILRLPLMIALHLG 729
Db 991 PFALDMDADLLIYVADLRDSE-----VAHLLQVCCYHLPLMOVELP- 1033
QY 730 RTHLNFQEFROONHLSCEFLVILGELLELOPHVFRSEHOGALMPCLLSFTILLNYRKRSSR 789
Db 1034 -----ISLTLRLALMP-----ISLNOFVNTVSASPR 1060
QY 790 HLAATFNKVFQITLHYTYNAPAAISFLQKHADPLDLSDNSDLVWLKSLGLSPSR 849
Db 1061 TIVSFLSVALLSDQPLTSLTSLLAHTARVLSPSH-LSE-----IQCLLAGSDSESR 1112
QY 850 DQDTGDLDEGESESSAGSLPVSVSLFTPLRAEMAPMKRLSRQYVEDLLEVLSDI 909
Db 1113 PLUSLIGHPE-----NSVRAHYTRLGLHLLQHSMALEGALQOSGSLSLLLGLGDK 1164
QY 910 DEMSRRRPETLSFTSTNLORLMSAEECRNLAF-SIALSRMONSPSIAAFPTPEYCL 968
Db 1165 DPVVR-----CSASFVGNATAYAGPLALAAA-VPSMTQL 1201
QY 969 GSDQFEV---VOTALRN-----LPEYALLCOEHAVALHRAFLVGYGM---DPSAJS 1017
Db 1202 GPDQAGIRRVASALGNLGEGLGEEELQCEVPORL-----EWACGDQDPNVK 1250
QY 1018 E---ALRLHME 1026

Db 1251 EAALIALRSLOQE 1263
RESULT 12
US-09-392-277-25
; Sequence 25, Application US/09392277A
; Patent No. 6451977
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage
; APPLICANT: Arnon Rosenthal
; TITLE OF INVENTION: Fused
; FILE REFERENCE: P1272R1P1
; CURRENT APPLICATION NUMBER: US/09/392,277A
; EARLIER APPLICATION NUMBER: US 09/258,000
; EARLIER FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: US 60/076,072
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 28
; SEQ ID NO 25
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence
US-09-392-277-25
Query Match 2.3%; Score 121; DB 4; Length 1315;
Best Local Similarity 20.0%; Pred. No. 0.025;
Matches 243; Conservative 130; Mismatches 426; Indels 414; Gaps 54;
QY 44 LVOTSEBALLPWLKRLMRSEVLRVDAALDLEPQQLLTVQSGFIPVSS-MSKLLQ 102
Db 235 LTPDPRKRLSWPDLINHPFAGHYTTITTEPAGPDL-----GTFPSTRPELO 282
QY 103 FLQDAVAH--DPTLEONINDKNYMAHLVEVQHERGASGGQTH----- 144
Db 283 VLADQGNHRLAPRAGNOSRILTLQAYKRMAEBAMQKHQNTGPALEQEDKTSKVAPTAPLP 342
QY 145 -----SL-----TASLP--PRDSTEAPKPKSSPEOPIGGRIR 177
Db 343 RLGATPOESSLAGILASSELKSSMAKSGTEGVSAPRENTTTPDCERAFPEE----- 394
QY 178 VGTOLRVIGED-DIAGMFQIIFPLSPDRMO-----SSSRPAL----- 217
Db 395 ---RPEVLGGRSDVD--LENEPDSDNEMOHLLETTEVPIQLKAPLTLICNPFCOR 449
QY 218 ---ALQALGOELARVVGSPPEVPGITVRVLOALATLSSPHGALVMSHRSHFLACPL 274
Db 450 IQSOLHAGGQILKILLEGASHI-----LPAFVLSLSSCSDSVALYSCREAGLPGLL 505
QY 275 LRQL--COYRCVPQDTEGSSLEFLKVLLOMLDMSPGVGGPLRAOLRLAQAAS----- 328
Db 506 LSLRHSQESNSLQOOSWYGT-FLQDLMAYIQ-----AVFACTFNLERQTSDSIQ 555
QY 329 ---AGRLSDVVRGL-----LRLAELAFRODDEVVSSTYRAV-----IATLRSG 370
Db 556 VFOEAMNLFIDLGLKLLAQPDSEQTLRRDSLKCFVLCAMGNSAISKAFYSSLLTT 615
QY 371 EOCSEV-----PDLISKVLOGLEVRSPHLE-----LITAFSATADAASPPAC 416
Db 616 QQVVLGDLHGLTVPOLPVHTPOGAPQVSPLEQSEDIGALISSALAICTAPVGLPDC 675
QY 417 -----KPY-----VVVSSLLQOE-----EE 431
Db 676 WDAKEQVCWMLANQLTEDSSQLRPSLISGLQHPIILCHLKLKLVYSCCLVEGCLRLQOE 735
QY 432 PLAGKPGADGSLLEAV-RIGPSSGILLVDMLEMDPEVSSCCDLOLRLEFSRRKG--G 488
Db 736 PLA-----LESIFMLIQCKVKVYDWEESTVTLVF-----LSLLVFLQNLPCG 779
QY 489 QAOVPSFRPYLLTLFTHOSSWPTLHOCIRVLLGKSRQO----- 526

```

Db 780 MEKGS---DVALTETH-SHVVSLVSAACLLGOLGOQVTFDLQPMEMAAATHALSAP 835
QY 527 ---RFDSSASIDFLMACI-----HVPRIWQGRDQRTPOKREELVLRVQGE 570
Db 836 AEVNLTPPGSGCFDGLILLLOLLEOGRASLRIDMSSMTVMVLRHRSVLR--PE 893
QY 571 LISLV--ELLIAEATRSQD-----GDPAACSLIQART--PLLSCCCGDESEVR 617
Db 894 EASAOESELSSPPEPDTLTISPOGMALLSLAMATEQEPOLCISLCHGSLIMS 953
QY 618 VTEHLSGCIQMGDSVIGRRCRDLLLOLYQRPRLRVPPVYLHSEGAASSVCK---- 673
Db 954 ILKHL-----LCPSPFLNOL-RQAPGSEFLPVVLE-----SVCQLLCE 990
QY 674 ---DDGILHREITLLADTDSRALENGADASMCRKLAVAHPL-LLLRHLPMTAALLHG 729
Db 991 PFALDMADLLIVYALDLRDE-----VAHLQYCCYHPLMQVELP- 1033
QY 730 RTHLNEQEPKQNHLSCEFLHVLGLLELLQPHVRSHEOGALMDCLLFTIRLLNVRKSSR 789
Db 1034 -----ISLTLRLALMDP-----TSLNCFVNTVSASPR 1060
QY 790 HLAFLNKVFQFIHKYTYNAPPAISFLQKHADPLHLSFSDNVLMLKSLLAGLSLPSR 849
Db 1061 TTVSFLSVALLSDQPLTSLDLSLAHTARVLSPSH-LSF-----IQELLASDESYSR 1112
QY 850 DDRTDRLDEGESESSAGSLPLVSVSLFTPLTAEMAPYKRLSRQTVDELVLSDI 909
Db 1113 PLRSLGHP-----NSVRANTYRLGLHLLQHSMAKRALOSQSLSLLLGLGDK 1164
QY 910 DEMKRREPELISFSTNLQRLMSAEBCRNLA-SLALNSMONSPSTAAAFLETFMYCL 968
Db 1165 DPVVR-----CSASFAGNANAYQAGPGLPAA--VPSMQLL 1201
QY 969 GSQDFEV---VOTALRN-----LPEYALLCOEHAIVLLHRAFLVGMVGOM---DPSAQIS 1017
Db 1202 GDPQAGIRRVNVAALGMLGPGGLGEBLLQCEVFORLL-----EMAGCPQPPNVK 1250
QY 1018 E---ALRILHME 1026
Db 1251 EALIALRSIQOE 1263

RESULT 13
US-08-952-127-12
Sequence 12, Application US/08952127
Patent No. 6211336
GENERAL INFORMATION:
APPLICANT: Shiloh, Yosef
APPLICANT: Tagle, Danilo A.
APPLICANT: Collins, Francis S.
TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6211336thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/952,127
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,995

```

```

REFERENCE/DOCKET NUMBER: 2290, 00029
TELECOMMUNICATION INFORMATION:
TELEPHONE: 810-539-5050
TELEFAX: 810-539-5055
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3066 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-952-127-12

Query Match
Best Local Similarity 18.4%; Pred. No. 0.14;
Matches 180; Conservative 136; Mismatches 254; Indels 410; Gaps 49;

QY 63 IRSEVRLVDAIADL-EPQDLLEFV-----QSFQIVSSMSKLLQFL----- 104
Db 925 IRRLLLLDSSIDLMKPLHLHMYLVLLKDLPGNEHSLPREDVVELQPLSLVCSLHRR 984
QY 105 DQVAHDPQTELEQIMDKNMAHLVEVQHERGASGGQTFHSLSLTLSPRRDSTEAKPK 164
Db 985 DQVVC---KTLISNVL-----HTV----- 1000
QY 165 SSPQPIGCG-----RIRVGTQLRVLPEDDLAGHFLQIFPLSPDRMOSSPRVAL 217
Db 1001 ---TNLQGSVDMESTRINAGHFLVWG-----AFHLTEKKECVF 1038
QY 218 ALQALCOEELARVYQGSB-----EVPQTVRVLAOLATLLSPHCGALVMSHRSHFL 270
Db 1039 SVRMALVCQLQTLLEADPYSEMAILNVKGFDPVNEAFSOFLADH-----HOVRML 1090
QY 271 ACPLRLQCYQRCVPODGTGSSFLKVLQMLQMLDSPVEGGPLRAO---LRLMASOA 327
Db 1091 AAGSVNRFLPQDMR---QGFSLSLAL-----BLFQOTSFNNATTTA 1130
QY 328 SAGRRISDVRCGLL-----RLAEALAFRODLEVYSSTV-----RAVITATLSGE 371
Db 1131 EAGIR-----GLLCDSONPDLDEIYNRKSVLLMIAVVLHCSPVCEKALFALCKSVK 1184
QY 372 QCSVEPDLISVLOGLEIVRPHLELLTAFFSATADASFPAKCKVYVSSLLQEBE 431
Db 1185 ENRLEPHLVKRV---LEKVSSEF-----GCR----- 1207
QY 432 PLAGCKPGADGSLFAVRLGPSSGLVDMLEMLDPEVYVSCPDQLRLILFSRRKGCQAQ 491
Db 1208 -----SLDEFMISHLDVLVLEMLNLODTER-----S 1233
QY 492 VPSFRPYLLTLEF-----HSSSWPTL--HOCIRVLLGKSREQRDPSPAS---LDPLMAC 540
Db 1234 LSSF-PEMLNVTSTIEDFYRSCYKILIPHLVIR-----SHRDEKVSIAINOIOCKWKS 1284
QY 541 IHV---PRI-----MGS-RQRTPOKREELVLRVQGS----- 568
Db 1285 LLYDCEPFLVHLIPYAVESTRDSYVQKR--ETATKVVDTLKGEDFLGKQIDQVFTSN 1342
QY 569 -PELISVELILAEATR-SODGD---TPACSLIQARLPLLSCCGD----- 611
Db 1343 LPETIV--VELLMTLHETADSDASASATRALCF-----SGDLDPANPPYF 1388
QY 612 -DESVRVTHELSCGCIQMGDSVIGRRCRDLLLOLYQRPRLRVPPV-----VLLHSE 664
Db 1389 PSVHTQATFVAVISNCHTKFKSIL-----EILSKIPDSYOKILLALCEQ 1432
QY 665 GAASSSVCKLDGL---HRTITLLADTSDSRALENGADASMCRKLAVAHPL----- 715
Db 1433 AAEFTNNVEKKRILIKIYHLEFSLCLKDIQS-----GLGGMAVAVLDVYITTLHYINKR 1486
QY 716 -----LRLHPMIALLHGRTHLNFOEPRQNHLSCEFLHVLGLLELLQPHVRSHEOG 768

```


Db 1487 SSHFTVSLRSFSLCCDLLSRVCHTAVTQCK--DALESHELV-IVGTILPLVDYOEW- 1541
Qy 769 ALMDCLSEFIRLLNLYKRSSRLIAFIN-----KFOQ 800
Db 1542 ---EYVLLKLYVDKNDKMLSTIKLIDPFPHVIFKDLRLOQIKYSGGFSLLIE 1598
Qy 801 FHKYTYNA--PAAIS-----FLOKHADPLHDL---SPDN-SDLYMLKSLAQL 844
Db 1599 EINHTLSAVNPLPLTRLEGKLDLRQLEQHKQDMDLRRASQNPQDGIYVKLVLSLL 1658
Qy 845 SLPSRDRTDRLGDEGEER 864
Db 1659 QL-----SKAVNQTGERE 1672

RESULT 14
US-09-651-656-15
; Sequence 15, Application US/09651656
; Patent No. 6340566
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES
; FILE REFERENCE: IL-10689
; CURRENT APPLICATION NUMBER: US/09/651,656
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 15
; LENGTH: 819
; TYPE: PRT
; ORGANISM: Thermus thermophilus
US-09-651-656-15

Query Match 2.3%; Score 119.5; DB 4; Length 819;
Best Local Similarity 22.8%; Pred. No. 0.015;
Matches 198; Conservative 81; Mismatches 296; Indels 293; Gaps 42;

Qy 17 PPRADSEFOALLDIWPEPEKPLPTAFIVDTSEALLPDMUKLRINSEVLRLVDALQ 76
Db 82 PIRADPATAERILKMGF-----RLAVALDQVEPAEEA-----EGLVREVQTL----- 124
Qy 77 DLEPOQLLFVOSFGIPVSSMSKLLQFLDQAVAHDPQLEQNM---DKNYAHVLEVOH 133
Db 125 -----TPGTLTQALLPREANVYLAIA----- 146
Qy 134 ERGASGGQTFHSLTASLPPRRDSTEARPKSSPEOPIGQR---IRVGTOQLRVLGPEDD 190
Db 147 -TGDWGLAFLDYSTGEF-----KGTILKSKSALYDELFRHRAPEVLLAPELR---ENEA 197
Qy 191 LAGMFLQITFP--LSPDPKMOSSPPRVALLAQALGQELARVVOGSPVEPGITTVRLQAL 248
Db 198 FVAERKRREPVALSEAP--FEPOGEGPLARRAAGALLAVARATOGG---ALSVRPER-- 250
Qy 249 ATLLSPHCGALVGMHSHFLACPLRLQCYORCVDPDTGSSFL-----KVLLOM 302
Db 251 ---LVDP--GAFVRLPEAS-----LKALEVEEPLRGDTLFGVYDETRTAPGRRLD- 297
Qy 303 LQWLSPGVEGGPLRAQQLMLASQASGRRLSDVVGGLRLLEALAFRODLEVSSTVA 362
Db 298 -AMLHPLLEKPLRARLDREVERFREGALRGVRRLLFRLLDLERLATRLLSRASPSD 356
Qy 363 VIATLRSEQGSVEPDLISKVLQGLI--EVRSPHLEELLTAFFSATADASPACPKPV 420
Db 357 LAALRRSLE-----ILPELKGILGEEVGLPDLISGLEELRAA----- 393
Qy 421 VVSSILLQEEPLAGKREGADGSLAVARLGGSSGLVDWMLDPEVYVSSCPDQLRL 480
Db 394 -----LVEDPPL-----KVSEGG---LIREGYDPLD--DALRRAHAEQVAYFLDEAR-- 436

Qy 481 ESRRKGOQOQVPSF-----RPY-----LITLFTHOS-SWPTLHO 514
Db 437 -----EKERTGPTLKVGNNVFGYYLETRVYEKQVQENRPVOTLMDROKTYLPEKKE 491
Qy 515 CIRVL-----LQKSREORPDS-----ASDPLMACIHP----- 544
Db 492 RERELVLEALIKRREEEFLLRBRARKEAALREARILAEVYALAEVAVRHGT 551
Qy 545 -----RIMGR-----DQRT-----PQKREELVIRVOGE-----LISIV 575
Db 552 RPFGEERLRIRAGRHPRVVERRTAEVNPDLMAHEVL--VTGNMAGKSTFLQTLIAL 610
Qy 576 ELI-----LAEATRSODS-----DTACSLIOA--RLPLLS 606
Db 611 AQISFVAEEAEELPLFDSITRIGASDOLAGKSTFWEMEVEVALYKEATERSIVLL- 669
Qy 607 CCCGDESRYKYTEHLG--CIGQMGDSVLRGRCDLLQTLQRLPRLRVPE----- 658
Db 670 -----DEVGRGTSSLDGVAIATALAEALHERRCYTLFATHYFELTALPLKMLHVA 723
Qy 659 -----VLLHS--EQAASS--VCKLDGLIHRTITLADTSRALENGADASMAC 705
Db 724 KEEGGLVFFHYOVLGPASKSYGEVYAEAGLPKEVERARALLSMARREGALEVLE 783
Qy 706 RKLAV-----AHPDLLRLPMTALLHG 729
Db 784 RLALDPRLTPLLEALRFLHEIKALALG 811

RESULT 15
US-09-650-855-15
; Sequence 15, Application US/09650855
; Patent No. 6365355
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA
; TITLE OF INVENTION: MISMATCHES
; FILE REFERENCE: IL-10284
; CURRENT APPLICATION NUMBER: US/09/650,855
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 15
; LENGTH: 819
; TYPE: PRT
; ORGANISM: Thermus thermophilus
US-09-650-855-15

Query Match 2.3%; Score 119.5; DB 4; Length 819;
Best Local Similarity 22.8%; Pred. No. 0.015;
Matches 198; Conservative 81; Mismatches 296; Indels 293; Gaps 42;

Qy 17 PPRADSEFOALLDIWPEPEKPLPTAFIVDTSEALLPDMUKLRINSEVLRLVDALQ 76
Db 82 PIRADPATAERILKMGF-----RLAVALDQVEPAEEA-----EGLVREVQTL----- 124
Qy 77 DLEPOQLLFVOSFGIPVSSMSKLLQFLDQAVAHDPQLEQNM---DKNYAHVLEVOH 133
Db 125 -----TPGTLTQALLPREANVYLAIA----- 146
Qy 134 ERGASGGQTFHSLTASLPPRRDSTEARPKSSPEOPIGQR---IRVGTOQLRVLGPEDD 190
Db 147 -TGDWGLAFLDYSTGEF-----KGTILKSKSALYDELFRHRAPEVLLAPELR---ENEA 197
Qy 191 LAGMFLQITFP--LSPDPKMOSSPPRVALLAQALGQELARVVOGSPVEPGITTVRLQAL 248
Db 198 FVAERKRREPVALSEAP--FEPOGEGPLARRAAGALLAVARATOGG---ALSVRPER-- 250

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 13, 2003, 23:29:27 ; Search time 91 Seconds
(Without alignments)
3467.806 Million cell updates/sec

Title: US-09-929-769-7

Perfect score: 5241
Sequence: 1 MHLLVYHAWMTLITLGPRA.....MDSQAQSEALRIILHFAVM 1029

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xjh
-Q/cgn2_1/USPIO.spool/US09929769/rn1-07052003.161925.8057/app.query.fasta.1.1223
-DB=issued.patents_NA -OPM=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09929769.qcgn.1.1.46.6runat_07052003.161925.8057 -NCP=6 -ICPU=3
-NO.XLPRY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:* Issued Patents.NA.*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCUTS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfill1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	157	3.0	33529	4	US-09-144-085-3
2	156.5	2.9	49377	1	US-08-764-233A-1
C 3	155.5	2.9	4080	2	US-08-446-345-35
4	150.5	2.9	28958	1	US-08-258-261B-6
5	150.5	2.9	28958	1	US-08-456-837-6
6	150.5	2.9	28958	1	US-08-457-342-6
7	150.5	2.9	28958	1	US-08-457-646A-6
8	150.5	2.9	28958	1	US-08-458-076A-6
9	150.5	2.9	28958	1	US-08-764-233A-4
10	150.5	2.9	28958	1	US-08-457-335A-6
11	150.5	2.9	28958	1	US-08-729-214-6
12	150.5	2.9	28958	3	US-09-028-934-6

13	147	2.8	4031	2	US-08-993-118-1	Sequence 1, Appl1
14	147	2.8	4031	3	US-08-845-528C-1	Sequence 1, Appl1
15	145	2.8	4225	3	US-08-993-118-9	Sequence 9, Appl1
16	145	2.8	4225	3	US-08-845-528C-9	Sequence 9, Appl1
17	145	2.8	4265	4	US-09-061-709-1	Sequence 1, Appl1
C 18	144.5	2.8	4154	1	US-08-131-365B-37	Sequence 37, Appl
C 19	144.5	2.8	4154	2	US-08-668-123-37	Sequence 37, Appl
20	135	2.6	8078	4	US-09-702-251-3	Sequence 3, Appl1
21	134.5	2.6	3879	2	US-08-916-352-1	Sequence 1, Appl1
22	134	2.6	5503	2	US-08-726-012B-1	Sequence 1, Appl1
23	133	2.5	3603	1	US-08-188-582-15	Sequence 15, Appl
24	133	2.5	3603	1	US-08-646-715-15	Sequence 15, Appl
25	132	2.5	3435	1	US-08-366-577-1	Sequence 1, Appl1
26	132	2.5	3435	1	PCT-US96-00005-1	Sequence 1, Appl1
C 27	131.5	2.5	36519	3	US-08-923-137-2	Sequence 2, Appl1
28	131.5	2.5	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
29	131.5	2.5	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
C 30	130	2.5	3490	2	US-08-841-483-3	Sequence 3, Appl1
C 31	130	2.5	3490	4	US-09-382-911-3	Sequence 3, Appl1
C 32	129.5	2.5	33529	4	US-09-144-085-3	Sequence 3, Appl1
C 33	129	2.5	4094	4	US-08-841-483-5	Sequence 5, Appl1
C 34	129	2.5	4094	4	US-09-382-911-5	Sequence 5, Appl1
C 35	129	2.5	7812	4	US-09-368-590-1	Sequence 1, Appl1
36	128.5	2.5	1371	2	US-08-910-731-1	Sequence 1, Appl1
37	128.5	2.5	1371	2	US-08-795-395-1	Sequence 1, Appl1
38	127	2.4	15378	3	US-08-785-420-1	Sequence 1, Appl1
39	126.5	2.4	6763	2	US-08-756-506-23	Sequence 23, Appl
40	126	2.4	2827	2	US-08-708-541A-25	Sequence 25, Appl
41	126	2.4	2827	4	US-08-940-968-24	Sequence 24, Appl
42	126	2.4	2827	4	US-08-940-968-25	Sequence 25, Appl
43	126	2.4	35081	2	US-08-752-760A-1	Sequence 1, Appl1
44	125.5	2.4	5228	2	US-09-428-711A-15	Sequence 15, Appl1
45	125	2.4	6344	4	US-08-843-417-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-144-085-3/c
Sequence 3, Application US/091444085
Patent No. 6280999
GENERAL INFORMATION:
APPLICANT: Gustafsson, Claes
APPLICANT: Betlach, Mary C.
APPLICANT: Ashley, Gary
APPLICANT: Julien, Bryan
APPLICANT: Ziemann, Rainer
TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
FILE REFERENCE: 30062-20020.20
CURRENT APPLICATION NUMBER: US/09/144.085
CURRENT FILING DATE: 1998-08-31
EARLIER APPLICATION NUMBER: 09/010.809
EARLIER FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 33529
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-09-144-085-3

Alignment Scores:

Pred. No.:	0.00168	Length:	33529
Score:	157.00	Matches:	209
Percent Similarity:	34.29%	Conservative:	112
Best Local Similarity:	22.33%	Mismatches:	337
Query Match:	3.00%	Indels:	281
DB:	4	Gaps:	41

US-09-929-769-7 (1-1029) x US-09-144-085-3 (1-33529)

Oy 150 SerLeuProIaGArGaRspSerThrGluAlaProLysProLysSerProGluCln 169

[illegible][illegible]

QY 807 ThrTYRAsnAlaProAlaAlaIleSerPheLeuGlnLysHisAlaAspProLeuHisasp 826
 Db 23677 ACC---TCGGCCCCCGCGCGC----- 23680
 QY 827 LeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeu 846
 Db 23660 ----- 23660
 QY 847 ProSerArgAspAspArgThrAspArgGlyLeuAspGlnGlyLeuGlnGlnSerSer 866
 Db 23659 CCGAGCCGACACACACACACACCTCCGCAAGCTCTCCGCGCAAGGAGGTACACACCGAGC 23600
 QY 867 AlaGlySerLeuProLeuValSer----- 874
 Db 23599 -----CAGGTGCCCCCTTACATCCGAGCTACCCGTCACACCGCTCGGAGCGTTTGCAC 23546
 QY 875 ValSerLeuPheThrProLeuThrAlaAlaGlnMetAlaProTyrMetLysArgLeuSer 894
 Db 23545 GCTTCCTCATACAGCAGCAGCTGCGCCGCGCTTCTGCTGCTTCTGCGCGCA-CGC 23487
 QY 895 ArgGlyGlnThrValGlnAspLeuLeuGlnValLeuSerAspIleAspGlnMetSerArg 914
 Db 23486 CGC-----GAGGTACGCGAG 23472
 QY 915 ArgArgProGlnIleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAla 934
 Db 23471 CAGCGCGCCGACGCGCGCGCGCAGCTCGTCAGCGCAGCTCGACGACTC----- 23444
 QY 935 GlnGlnCysArgAsnLeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerPro 954
 Db 23423 -----CGCAC-----GCCATCTCCCGCCCTCCACGCGCGC----- 23391
 QY 955 SerIleAlaIleAlaPheLeuProThrPheMetTyrCysLeuGlySerGlnAspPheGlu 974
 Db 23390 -----ATTCCACAGCGCTG-----CTCCGCGCTGAGACATTTCAC 23355
 QY 975 ValValGlnThrAlaLeuArgAsnLeuProGlnTyrAlaLeu 988
 Db 23354 CTCGGTCCGCGGCTCTTCGCGGCTTCACAGCCAGTAGGCGCTG 23313
 RESULT 2
 US-08-764-233A-1
 : Sequence 1, Application US/08764233A
 : Patent No. 5716849
 : GENERAL INFORMATION:
 : APPLICANT: Ligon, James M.
 : APPLICANT: Schupp, Thomas
 : APPLICANT: Beck, James J.
 : APPLICANT: Hill, Dwight S.
 : APPLICANT: Neff, Suzanna
 : APPLICANT: Ryals, John A.
 : TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
 : NUMBER OF SEQUENCES: 10
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Ciba-Geigy Corporation
 : STREET: 520 White Plains Road, P.O. Box 2005
 : CITY: Tarrytown
 : STATE: NY
 : COUNTRY: USA
 : ZIP: 10591
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/764,233A
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/729,214
 : FILING DATE: 09-OCT-1996

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/258,261
 FILING DATE: 08-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Weigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: 1506/CIP6
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919) 541-8587
 TELEFAX: (919) 541-8689
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 49377 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Sorangium cellulosum
 IMMEDIATE SOURCE:
 CLONE: p98/1, pL3, and pVKM15
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 383..760
 OTHER INFORMATION: /product= "SorR"
 OTHER INFORMATION: /note= "This gene encodes a protein that is highly homolog
 OTHER INFORMATION: the reductase domains of type I PKs such as eryA from
 OTHER INFORMATION: Saccharopolyspora erythraea."
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 927..19874
 OTHER INFORMATION: /product= "Sora"
 OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs t
 OTHER INFORMATION: are known to be involved in the synthesis of polyketide
 OTHER INFORMATION: compounds."
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 942..7115
 OTHER INFORMATION: /product= "Module 1 of Sora"
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 7203..12884
 OTHER INFORMATION: /product= "Module 2 of Sora"
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 13455..19616
 OTHER INFORMATION: /product= "Module 3 of Sora"
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 19871..46318
 OTHER INFORMATION: /product= "Sorb"
 OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs ge
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 19870..24556
 OTHER INFORMATION: /product= "Module 1 of SorB"
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 24638..30820
 OTHER INFORMATION: /product= "Module 2 of SorB"
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 30881..35446
 OTHER INFORMATION: /product= "Module 3 of SorB"
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 35528..40114
 OTHER INFORMATION: /product= "Module 4 of SorB"
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: 40190..46318
 OTHER INFORMATION: /product= "Module 5 of SorB"
 FEATURE:

NAME/KEY: misc_feature
 LOCATION: 46851..47891
 OTHER INFORMATION: /product= "SORM"
 OTHER INFORMATION: /note= "The protein encoded by the sorm gene is highly homologous to the methyltransferase from Streptomyces
 OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the OTHER INFORMATION: polyketide rapamycin."

Alignment Scores:

Pred. No.: 0.00387 Length: 49377
 Score: 156.50 Matches: 274
 Percent Similarity: 35.038 Conservative: 119
 Best Local Similarity: 24.428 Mismatches: 435
 Query Match: 2.99% Indels: 301
 Gaps: 61

US-09-929-769-7 (1-1029) x US-08-764-233A-1 (1-49377)

QY 14 ThrLeuGlyProProAlaAlaAspSerGluPheGlnAlaLeuLeuAsp----- 30
 DB 42525 ACCTCCGGCAACCGCTCTGTC-TCGAGCGCGACGAGCGGCTCTCGACGATGGCAT 42583
 QY 31 IleTyrPheProGluGluGlyProLeuPro-----ThrAlaPheLeuValAspThrSer 48
 DB 42584 CGCTTCTGTCGAGGTACACCCCATCCGTCCTCAGCTGCGCCCTCCGCGAGACTGC 42643
 QY 49 GluGlnAlaLeuLeuLeuProAspTyrLeu---LysLeuArgMetIleArgSerGluVal 67
 DB 42644 GAGCGGTACACCGCTCATCCGTCGTCGTCGTCATTCGACGCGACGAGGACCACTTC 42703
 QY 68 LeuArgLeuValAlaAlaLeuGlnAspLeuGluProGlnGlnLeuLeuPheVal 87
 DB 42704 GCCCGCTGCTC---CTCTCTGCGGAGCTCTCTACCCGAGCGCTCGCGCTCGACTGG 42760
 QY 88 GluSerPheGlyIleProValAspSerMetSerLysLeuLeuGlnPheLeuAspGlnAla 107
 DB 42761 AACGCTTCTGCGCGCTTCTGCTCCCGCAAGTCTCTCCCACTTC----- 42811
 QY 108 ValAlaAlaAspProGlnThrLeuGlnGlnAsnIleMetAspLysAsnTyrMetAlaHis 127
 DB 42812 -----CCCTCCCAACGAGCGCTCTGCTGTCAGCGCTCC---ACGGCGCAC 42856
 QY 128 LeuValGluValGlnHisGluArgGlyAlaSerGlyGlyGlnThrPhe-----HisSer 145
 DB 42857 GCTGCGAGCTC-----GCCCTCGCAGGCTGACCTCGGCGGACCAACCGC 42901
 QY 146 LeuLeuThrAlaSerLeuProProArgArgAsp----- 156
 DB 42902 CTGCTGCGGCGCGCGCTGC-CCTCGCGACCGCGATGCTTGTCTTCACAGACGGCT 42960
 QY 157 -----SerThr 158
 DB 42961 CTCCTCGCAGACACCGGCTCGAAGACACAGTCGTTCGGCATACCTGCTCTCC 43020
 QY 159 GluAlaProLysProLysSerSerProGlu---GlnProIleGlyGlnGlyArgIleArg 177
 DB 43021 CAGCGCCCGCTCTCGAGCTCGCCTGCATGTGCGCCCATCTGCTGCGCTCGACACCGT 43080
 QY 178 ValGlyThrGlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGln 197
 DB 43081 CGAAGACGTACCGCTCGA----- 43098
 QY 198 IlePheProLeuSerProAspProArgTyrGlnSerSerSerProArgProValAlaLeu 217
 DB 43099 -----CCCCCGCTGCTCTCCATCGACGAGGCGCGCTCTCTCCCA--- 43140
 QY 218 AlaLeuGlnGlnAlaLeuGlyGlnGlnLeuAlaArgValAlaGlnGly---SerProGlu 226
 DB 43141 GATCTCTGTCGCGCGCGCGACGGT-CTGGAC-----GAGGCGCGCTCTCGCTTC 43190
 QY 237 ValProGlyIlePheValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHis 256
 DB ----- 44117

DB 43191 ATAGCGCGCGCACGACCGCTTC-----AGATAGCGCGCTGACTCGCCACGCCA--- 43241
 QY 257 GlyGlyAlaLeuValMetSerMetHisArgSerHisPheLeuAlaCys----- 272
 DB 43242 ---CGCGCTCTTCGCGCAAGCTACGCCGTC-CATTGCTTGTGATGCTCGCGAATGCG 43297
 QY 273 ---ProLeuLeuArgGlnLeuGlyGlnTyrGlnArgCysValProGlnAspThr---Gly 290
 DB 43298 CCCCC-----TCGGCGCGCACCGAGTGCACACCAAGGT 43333
 QY 291 PheSerSerLeuPheLeuValLeuLeuGlnMetLeuGlnIlePheAspSerProGly 310
 DB 43334 TTCTAGCAGCC-----CTCGAGCGCTGCG 43360
 QY 311 ValGluGlyGlyPro-----LeuArgAlaGlnLeuArgMet 322
 DB 43361 CTTCGTTATGCGCCCGAGTTCACAGGGCTCGCGCGCTCTTACAGCGCGCGACGAGT 43420
 QY 323 LeuAlaSerGlnAlaSerAlaGlyArgArg-----LeuSerAsp 335
 DB 43421 CTTCGCGGAGCCAGCTCCCGGACCGCGGAGAGAGACGCGCTGTTTCCCTCCA 43480
 QY 336 ValArgGlyGlyLeuLeuArgLeuAlaGlnAlaLeuAlaPheArg----- 350
 DB 43481 CCGCGCGCTGCTGACAGCGCTTGCAGCGCTCGCTTGTAGACGACCGCAAGGCAAGC 43540
 QY 351 ---GlnAspLeuGlnValAlaSerSerThrValArgAlaValIleAlaThrLeuArgSer 369
 DB 43541 CTTCAGAGATCCCTTCGCGGAGCGAGT-----ATCGGTGCGCTCGGCTG 43588
 QY 370 GlyGluGlnCysSerValGluProAsp-----LeuIleSerLysValLeuGlnGly 386
 DB 43589 -----GAGCACCAACCTGCGCGCTGCTTCACCGCTCGAGGCG 43630
 QY 387 LeuIleGluValArgSerProHisLeuGlnGluLeuThrAlaPhePheSerAlaThr 406
 DB 43631 AATCTCGCGCTGCTTC-----TCCTCGCGC 43657
 QY 407 AlaAspAlaAlaSerProPheProAlaCysLys-----ProValValValAlaSerSer 424
 DB 43658 ACCCGAGAGCGCAACCATGCTCGGTGCACAGCGCTCGCAATGCGCG----- 43705
 QY 425 LeuLeuLeuGlnGluGluProLeuAlaGlyGlyLysProGlyAlaAspGlyLys 444
 DB 43706 -----CCGCGTCGCGGAGAGCTCGCGACACCGGAGCG-----TCCAC 43747
 QY 445 LeuGluAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTyrPheGluMetLeu 464
 DB 43748 CTGATGCTCTC-----TTCGCGATGACGTGACGAGCTGCA- 43785
 QY 465 AspProGluValAlaSer-----SerCysAspAspLeuGlnLeuArgLeuLeu 480
 DB 43786 AAGCCCACTCACCGCCCTCCGCGGAGCGCGGCTGCTCGGACACAGAGG----- 43839
 QY 481 PheSerArgArgGlyGlyGlyGlnAlaGlnValProSerPheArgProTyrLeuLeu 500
 DB 43840 ---TCTGACCTGCGGACAGGCT---GCTCTGACACCTGATACGACCTTCTGCTCT 43893
 QY 501 Thr-----LeuPheThrHisGlnSerSerTyrProThrLeuHisGln 514
 DB 43894 ACGCAGCGCGCTGACAGGCGGCTTGCCTCAAGGCTGTATGCGCCCTTCATCGC 43953
 QY 515 CysIleArgValLeuLeuGlyLysSerArgGluGln-----ArgPheAspPro 530
 DB 43954 TCTGCGGAGGCGACCTTCGCGAGCGCGCGGACACCGACCGACCGCTGCGCT 44013
 QY 531 SerAlaSerLeuAspPheLeuThrAlaCysIleHisValProArgIleTyrGlnGlyArg 550
 DB 44014 CTTCGACAGCTG-----GCTCGCGGACGAGCGCTGCTCTGCGCGCTGCG 44061
 QY 551 AspGlnArgThrProGlnLysArgArgGluGlnLeuValLeuArgValGlnGlyProGlu 570
 DB 44062 ---CTGCTGACCGCGAGCGCGCTGCGCACCGCTGAGAGAGAGCTGACAGGCGCTCG- 44117

QY 571 LeuileSerLeuValGluLeuLeuLeuValGluValThrArgSerGlnAspLeu 590
 DB 44118 CTCACGGCCCTCTGGGGTCTGCTGCTCCGGCAGAGGAGCACCACAGCCCTC 44177
 QY 591 ThrAlaIaCysSerLeuLeuLeuAlaArgLeuProLeuLeuSerCysCysGly 610
 DB 44178 TCGTCCTGTCGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 44237
 QY 611 AspArgLeuSerValArgValThrGlnHisLeuSerGlyCysIleGlnGlnPro 630
 DB 44238 ACGCAGAGGAGCAGAGATCCCTCCGCAAGGCAACCCCTGTCCTCA 44291
 QY 631 AspSerValLeuGlnArgArgCysArgAspLeuLeuGlnLeuTyLeuGlnArgPro 650
 DB 44292 TCTCAGCGCTCCCGAGGCGCCAGG-----ACACAGCTGCC 44330
 QY 651 GluLeuArgValProValProGluValLeuHisSerGlnGlyAlaAlaSerSer 670
 DB 44331 CCGCAGGCTCGGAGGAGCAGCTC-CTCATCAGGAGGAGCAGCAGCTCGGCGCCTG 44389
 QY 671 ValCysLeuAspGlyLeuIleHisArgPheIleThrLeu 685
 DB 44390 GTCCGGCGCGCTGCTGTAACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 44449
 QY 686 -----AlaAspThrSerAspSerArgAlaLeuGlnAsnArgGlyAla 699
 DB 44450 GCGCGAGCGCTCGGCGCTGATGCTCTGGAGAGGAG-----CTGAGAGCTCTGGGGCT 44506
 QY 700 AspAlaSerMetAlaCysArgValLeuAlaValAlaHisProLeuLeu 718
 DB 44507 TCGGTACCTCGCGCTGCGAGTGCGCATCCAGCAGCTGCTCAAGAGCTTGTGAT 44566
 QY 719 HisLeuProMet-----IleAlaAlaLeuHisGlyArgThrHisLeuAsn 735
 DB 44567 AACATTCGAGCGCTCACCCGCGCGCGCTGCGCATGCGCGCGCTGCGAGCG 44626
 QY 736 GlnGluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGly 755
 DB 44627 GAT-----CTGCTGCGCGCGCATGAC 4467
 QY 756 LeuLeuGln-----ProHisValPheArgSerGlnHisGlnGlyAlaLeuThrAspCysLeu 774
 DB 44648 CTCGAGCGGATGACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 44698
 QY 775 LeuSerPheIleArgLeuLeuAsnTyArgLysSerSerArgHisLeuAlaPhe 794
 DB 44699 -----TTGCATCAGCTCACCCAGATTAACCCCTTGCCTGCGCTTC 44737
 QY 795 Ile-----AsnLysPheValGlnPheIleHisLysTyIleThrTyAsnAlaProAla 812
 DB 44738 ATCCCTTCTGTCGCGCGCGCGCGCTCTCGGACGCTCAGTCACTCACTAGCCGCT 44797
 QY 813 AlaIleSerPheLeuGlnLysHisAlaAsp-----ProLeuHis 825
 DB 44798 GCGAGCGCTTCTGATGCGCTTGCACACCGCGCGCGCGCGCGCGCGCGCTCTCA 44857
 QY 826 AspLeuSerPheAspAsn-----SerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeu 844
 DB 44858 TCGCTCGGTCGAGCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 44917
 QY 845 SerLeuProSerArgAspArgThrAspArg-----GlyLeuAspGlnGlnGlyValGln 863
 DB 44918 GCGCGCGCT-----CGCATGAGCGCGCGCGCGCTTCTCCCTCAGCTCTGAGGAG 44965
 QY 864 GluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThr 882
 DB 44966 AGG-----CTGCGCGCTTCTGATGCGCGCTCTTCTCCAGCGAGCGCGCG 45010
 QY 883 -----AlaAlaGlnMetAlaPro 888
 DB 45011 CTGGTCCCGCGCGCTTCTGACTTGAAGCGCGCTCAAGGCGGAGCGCGCGCGCGCG 45070

QY 889 TyrMetLysArgLeuSerArgGlyGlnThrVal----- 899
 DB 45071 TTGTTCCACAGCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 45130
 QY 900 --GluAspLeuLeuGlnValLeuSerAspIleAspGlnMetSerArgArgProGln 918
 DB 45131 GCCTGCTGCTTACAGAGCGCGCTCTCAGCCTCCCGCGCGCGCGCGCGCGCGCG 45187
 QY 919 IleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGln---GlnCys 937
 DB 45188 CTGCTGATCTCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 45247
 QY 938 CysArg-AsnLeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAl 957
 DB 45248 CTCGATCCCGATTCGCTCTCTCCAGAGCTTGCGCTGAGCTCATCATCAGCTGATCTC 45307
 QY 957 AlaAlaPheLeuProThrPheMetTyCysLeuGlySerGln-----AspPheGluVal 975
 DB 45308 CGAATCGGCTCG 45364
 QY 975 IValGlnThrAlaLeuArgAsnLeuPro-----GluTyAlaLeuLeuCysGln 991
 DB 45365 CATCAACCCAGCTGCGCTGCGAGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 45424
 QY 991 nGlu 992
 DB 45425 AGAG 45428

RESULT 3
 US-08-446-345-35/c
 : Sequence 35, Application US/08446345
 : Patent No. 5831009
 :
 : GENERAL INFORMATION:
 : APPLICANT: WILRICH, Axel
 : APPLICANT: MOLLER, Niels P.H.
 : APPLICANT: MOLLER, Karin B.
 : TITLE OF INVENTION: NOVEL PROTEIN PHOSPHOTYROSINE
 : TITLE OF INVENTION: PHOSPHATASES PTP-DI
 : NUMBER OF SEQUENCES: 41
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: N.Y.
 : COUNTRY: U.S.A.
 : ZIP: 10036-2711
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/446,345
 : FILING DATE: 22-MAY-1995
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/234,440
 : FILING DATE: 28-APR-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Coruzzi, Laura A.
 : REGISTRATION NUMBER: 30742
 : REFERENCE/DOCKET NUMBER: 7683-054
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (212)790-9090
 : TELEFAX: (212) 869-8864
 :
 : INFORMATION FOR SEQ ID NO: 35:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4080 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: both
 : TOPOLOGY: unknown
 : MOLECULE TYPE: CDNA


```

Db 28435 ATCTCTTCTGTCGTCGCGCGCTCCTCGGACAGTCAGTCACCACTACGCCGT 28494
Qy 813 AAlaIleSerPheLeuGlnLysHisAlaAsp-----ProleuHis 825
Db 28495 GCGAGCGGCTTCCTGATGCGCTTCGACACCGCGCGCGCGCAAGGCGCTCCCTCA 28554
Qy 826 AspleuSerPheAspAsn---SeraspLeuValMetLeuLysSerLeuLeuAlaGlyLeu 844
Db 28555 TCGCTCGGCTGGAGCAGCTGGCGCCGACGCGCAGCGCATGACAGACAGCTGACGCCGCC 28614
Qy 845 SerleuProSerArgAspAspArgThrAspArg---GlyLeuAspGlnGlnGlyLeu 863
Db 28615 GCGCGCCCT-----CGCATGGAGCGCGCGCGCTTCCTCGACCTGTGAGGAG 28662
Qy 864 GluSerSerAlaGlySerleuProLeuValSerValSerLeuPheThrProleuThr--- 882
Db 28663 AGG-----CTCGCCCTTCGATGGCGGCTTCGCAACCGAGCCGCC 28707
Qy 883 -----AlaAlaGlnMetAlaPro 888
Db 28708 CTGATCCCGCGCGCTTCGACTTGAAGCGCGCTCAGGCGGACGCCGCGCTCCCGCG 28767
Qy 889 TyrMetLysArgLeuSerArgGlyGlnThrVal----- 899
Db 28768 TTGTTCCAAAGCTGCTCGCGCTCGACCGCAAGCGCGCGCGCAGACACCGCCGAG 28827
Qy 900 ---GluAspLeuLeuGlnValLeuSerAspIleAspGlnMetSerArgArgProGln 918
Db 28828 GCGTCGTCGCTTACAGACGCGCTTCACGCTCCCGCGCGCGCGAGAGGT---GCC 28884
Qy 919 IleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAla 934
Db 28885 CTGCTGCATTCATCCGACCGACCGCGCGCGCTTCGCGCTCGCC 28932

```

RESULT 5

```

US-08-456-837-6
Sequence 6, Application US/08456837
Patent No. 5643774

```

GENERAL INFORMATION:

```

APPLICANT: Schnupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryalls, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF INVENTIONS: antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,837
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:

```

```

NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
FAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-456-837-6

```

Alignment Scores:

```

Pred. No.: 0.006 Length: 28958
Score: 150.50 Matches: 255
Percent Similarity: 34.668 Conservative: 111
Best Local Similarity: 24.158 Mismatches: 405
Query Match: 2.87% Indels: 292
DB: 1 Gaps: 57

```

US-09-929-769-7 (1-1029) x US-08-456-837-6 (1-28958)

```

Qy 14 ThrLeuGlyProProArgAlaAspAspSerGluPheGlnAlaLeuLeuAsp----- 30
Db 26222 ACCCTCGGCAACCGCTCTGCTTC-TGAGCGGACCGCGCGCTCCTCGACGATGGCAT 26280
Qy 31 IleTrpPheProGlnGlnLysProLeuPro-----ThrAlaPheLeuValAspThrSer 48
Db 26281 CGCTTCTTGAGAGTACAGGCGCCCATCCGTCGTCAGAGCTCGCCCTCGCGACGCTCG 26340
Qy 49 GlnGlnAlaLeuLeuLeuProAspTrpLeu---LysLeuArgMetLysSerGlnVal 67
Db 26341 GAGCGCTACCGCTCGATCCGCTCGTCGCGCTCATTCAGACGAGGACGACGACCTTC 26400
Qy 68 LeuArgLeuValAspAlaAlaLeuGlnAspLeuGlnProGlnGlnLeuLeuLeuVal 87
Db 26401 GCCCGCTGCTC---CTCTCGGCGGAGCTCTTACCGGAGGCTCGCGCTGCACTGG 26457
Qy 88 GlnSerPheGlyIleProValSerMetSerLysLeuLeuGlnPheLeuAspGlnAla 107
Db 26458 AACGCTTCTTCGCGCGCTTCGCTCCCGCAAGGCTCCCTCCCACTAC----- 26508
Qy 108 ValAlaHisAspProGlnThrLeuGlnGlnAsnIleMetAspLysAsnTrpMetAlaHis 127
Db 26509 -----CCCTTCACAGCGGAGCGCTTCGCTGCGTGCAGCGCTCC---ACGCGGAC 26553
Qy 128 LeuValGlnValGlnHisGlnArgGlyAlaSerGlyGlnThrPhe-----HisSer 145
Db 26554 GCTGCGGAGCTC-----GCCCTCGGAGGCTGACCTCGCGCGACGACCGCG 26598
Qy 146 LeuLeuThrAlaSerLeuProProArgArgAsp----- 156
Db 26599 CTGCTCGGCGCGCGCGCTGC-CCTGCGGACCGCGCATGCTTGTCTTCACAGAGCGCT 26657
Qy 157 -----SerThr 158
Db 26658 CTCCTCGGAGACACCGCTGCTCGAAGACGACGCTTCGCGATACCTGTCTGTC 26717
Qy 159 GluAlaProLysProLysSerSerProGlu---GlnProIleGlyGlnGlyArgLeu 177
Db 26718 CAGGCGCGCGCTTCGACGCTGCGCTGCATGCTCGCCATCTCGCGCTGACACGCT 26777
Qy 178 ValGlyThrGlnLeuArgValLeuGlyProGlnAspAspLeuAlaGlyMetPheLeuGln 197
Db 26778 CGAAGACGTCAGCTCGA----- 26795
Qy 198 IlePheProLeuSerProAspProArgTrpGlnSerSerSerProArgProValAlaLeu 217

```

Db 26796 -----CCCCCCTCGCTTCACATCGACGCGCGCTCTCTCCA--- 26837
 QY 218 AlaLeuGlnAlaLeuGlnGluLeuAlaArgValValGlnGly---SerProGlu 236
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 26838 GATCTCGTGGGGCCGCGGACGGTG-CTGGAC-----GAGGCGCCTCTCGCTTC 26887
 QY 237 ValProGlyIleThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerProHis 256
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 26888 ATGACGGGCGCCACGACGCGCTTC-----AGATGCGCCCTGACCTCGCCACGCCA--- 26938
 QY 257 GlyIleValLeuValMetSerMetHisArgSerHisPheLeuLacys----- 272
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 26939 ---GCGGCTCTCTCGCCCAAGCTACGCCGTCC-CATTGCTTGAGTCTCCGCGATGCC 26994
 QY 273 ---ProLeuArgGlnLeuGlnGlyGlnTyrGlnArgCysValProGlnAspThr---Gly 290
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 26995 CCCCC-----TCGGGCGCCACCCAGGAGGACCCAGAGT 27030
 QY 291 PheSerSerLeuPheLeuLysValLeuLeuGlnMetLeuGlnTrpLeuAspSerProGly 310
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27031 TTCTACGACAGCC-----CTCGAGAGCGCTGCG 27057
 QY 311 ValGlnGlyIlePro-----LeuArgAlaGlnLeuArgMet 322
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27058 CTTCGCTTATGGCCCGCATTCACAGGCGCTCCGCCGCTCTACAGCGCGCAGAGCT 27117
 QY 323 LeuAlaSerGlnAlaSerAlaGlyArgArg-----LeuSerAsp 335
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27118 CTTCGCGCGAAGCCAGAGCTCCGCGAGCGCGCGCAAGAGAGCGCGCTTTGCGCTCCA 27177
 QY 336 ValArgGlyIleLeuLeuArgLeuAlaGlnAlaLeuAlaPheArg----- 350
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27178 CCGCGCGCTCTCGACAGCGCTTCGAGCGCGCTGCTTGTAGACAGCAGCAGGAAAGGC 27237
 QY 351 ---GlnAspLeuGlnValAlaSerSerThrValArgAlaValIleAlaThrLeuArgSer 369
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27238 CTTCAGAGATCCCTTCTCGTGGAGCGAGT-----ATCGCTGGCGCTCGGCTG 27285
 QY 370 GlyIleGlnCysSerValGluProAsp-----LeuIleSerLysValLeuGlnGly 386
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27286 -----GAGCCACACCCCTGGCGCGCTTCACCGCTTCAGCGCTGAGGCG 27327
 QY 387 LeuIleGlnValArgSerProHisLeuGlnGluLeuLeuThrAlaPhePheSerAlaThr 406
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27328 AATCTCGCGCTCGCTCC-----TCTCTCCCG 27354
 QY 407 AlaAspAlaAlaSerProPheProAlaCysLys-----ProValValAlaValSerSer 424
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27355 ACGCCAGAGCGAAGCCATCGCTCGGTGCAAGCGCTCGCATCGCGC----- 27402
 QY 425 LeuLeuLeuGlnGluGluProLeuAlaGlyIleLysProGlyAlaAspGlyLys 444
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27403 -----CCGCGCTCGCGAGAGCTCGCGAGACCGCGGAGCG-----TCCAC 27444
 QY 445 LeuGlnAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrpLeuGlnMetLeu 464
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27445 CTGGAATCCCTC-----TTCGCGATCGACCTGAGACGACTGCA- 27482
 QY 465 AspProGluValAlaSer-----SerCysProAspLeuGlnLeuArgLeuLeu 480
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27483 AAGCCCACTACACCGCGCATTCGCCCGAGCGGCTCGCTCGGCACAGAGS----- 27536
 QY 481 PheSerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTyrLeuLeu 500
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27537 ---TCTGACCTCGGAGCAGAGGT---GCCTTCGACCGCTATACGACCTGTGTGCTCT 27590
 QY 501 Thr-----LeuPheThrHisGlnSerSerTrpProThrLeuHisGln 514
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27591 ACGCAGCGCCCTCGACAGGCGCTTCCTCCACAGCCTCGATCGCCCTTCATACGC 27650
 QY 515 CysIleArgValLeuLeuGlnGlyLysSerArgGln-----ArgPheAspPro 530
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27651 TCTGCCGAGAGGAGACTCATCTCGAGAGCGCCCGAGACACCGCGAGCTCGCTCT 27710

QY 531 SerAlaSerLeuAspPheLeuTrpAlaCysIleHisValProArgIleTrpGlnGlyArg 550
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27711 CTTCGAAGCCTG-----GCTGCGCGAGAGAGCGCTCGCTCTCGCGCGCTCG 27758
 QY 551 AspGlnArgThrProGlnLysArgArgGluGluLeuValLeuArgValGlnGlyProGlu 570
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27759 ---CCTGCTACCGAGCGCGCGCTCGCCACCGCATGGAAGACCTCAAGGCGCTCG- 27814
 QY 571 LeuIleSerLeuValGluLeuIleLeuAlaGlnAlaGlnThrArgSerGlnAspGlyAsp 590
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27815 CTCAGCGCGCTCTCTGGGGCTCGCTCGCTCGCCGCGAGAGACAGACACCGAGCGCTC 27874
 QY 591 ThrAlaAlaCysSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCysGly 610
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27875 TCGTCTCGTGCAGCTCGACAGACAGAGCGAGCGCTCCAGCAGCGCTCTCGGCGCTCG 27934
 QY 611 AspAspGlnSerValArgLysValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGly 630
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27935 ACGCAGAGAGAGCCAGAGATGCGCTCGCAAGCGCAACCCCTCGTTCAA-----GGC 27988
 QY 631 AspSerValLeuGlyArgArgCysArgAspLeuLeuGlnLeuLysTrpLeuGlnArgPro 650
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 27989 TCTACGCTGCGCCCGAGCGCGCCAGCG-----ACACAGCGCTCC 28027
 QY 651 GluLeuArgValProValProGluValLeuLeuHisSerGlnGlyAlaAlaSerSerSer 670
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28028 CCGAGAGCTCTCGAGGACCGCTC-CTCATCAGGAGAGCAGCGGACGCTCGCGCGCTG 28086
 QY 671 ValCysLysLeuAspGlyLeuIleHisArgPheIleThrLeu----- 685
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28087 GTGCGCGCGCGCTCGTGAACACAGAGCGCAAGCACCTGCTCTCACTCGCGCGCG 28146
 QY 686 -----AlaAspThrSerAspSerArgAlaLeuGlnAlaAspArgGlyAla 699
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28147 GCGCGAGCGCTCGCGGTGCTGATGCTCTGCGAAGCAG-----CTCGAAGCTTGGGGCT 28203
 QY 700 AspAlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeu---LeuLeuArg 718
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28204 TCGGTACCTCGCGCGGTGCGAGCTGCGCATCCAGCAGCGCTCAAGGAGACTTGTGAT 28263
 QY 719 HisLeuProMet-----IleAlaAlaLeuLeuHisGlyArgThrHisLeuAspPhe 735
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28264 AACATTCGAGCGCTCACCGCGGTGCGCTGCGTGTGATCCCGCAGCGCTCGAGAGGC 28323
 QY 736 GlnGluPheArgGlnGlnAlaHisLeuSerCysPheLeuHisValLeuGlyLeuGlu 755
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28324 GAT-----CTGCTCGCGCGCATGAGC 28344
 QY 756 LeuLeuGln---ProHisValPheArgSerGluHisGlnGlyAlaLeuTrpAspCysLeu 774
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28345 CTGAGCGGAGTACGACCGCGCTTCGCCCAAGATCGATCCCGC---TGGCAC----- 28395
 QY 775 LeuSerPheIleArgLeuLeuLeuAsnTyrArgLysSerSerArgHisAlaAlaPhe 794
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28396 -----TTCATCAGCTACCCCAATATAGGCCCTTCCGCGCTTC 28434
 QY 795 Ile-----AsnLysPheValGlnPheIleHisLysTyrIleThrTyrAsnAlaProAla 812
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28435 ATCCTCTTCGTCGTCGCGCGCGGTCTCGGAGCTCAGGTACATCAGTACAGCGCT 28494
 QY 813 AlaIleSerPheLeuGlnLysHisAlaAsp-----ProLeuHis 825
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28495 GCGAGCGCTCTCTCGATGCGCTTGGCGCACACCGCGCGCGAGAGGCTCTGCTGCTCA 28554
 QY 826 AspLeuSerPheAspAsn---SerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeu 844
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28555 TCGTCTCGTGGAGACCTGCGGAGCGAGCGCAATGACAGACAGCTACAGCGCGCC 28614
 QY 845 SerLeuProSerArgAspAspArgThrAspArg---GlyLeuAspGlnGlnGlyGln 863
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
 Db 28615 GCGCGCGCT-----CCATGAGAGCGCGCGCTTCCCTCGAGCTCGAGAGAG 28662

QY 864 GluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThr 882
 Db 28663 AGG-----CTGGCCCTCTTCGATGCGGGGCTCTCCAGACCGAGCCGCC 28707
 QY 883 -----AAlaAlaGluMetAlaPro 888
 Db 28708 CTGGTCCCGCGGCTTCGACTTGAGCGCGCTCAGGCGGAGACGCCGCGAGCTCCCGC 28767
 QY 889 TyrMetLysArgLeuSerArgGlyGlnThrVal----- 899
 Db 28768 TTGTTCCACGCTCTGCTCCGCGCTCCAGCCGACGACGAGCCGCCAGACACCGCCAG 28827
 QY 900 ---GluAspLeuLeuGluValLeuSerAspLeuAspGluMetSerArgArgArgProGlu 918
 Db 28828 GCGTCGTCGCTTACAGAGCGGCTCTCAGCCCTCCCGCGCGCGGAGAGCGGAGCGT---GCC 28884
 QY 919 IleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAla 934
 Db 28885 CTGCTCGATCTCATCCGACCGAGCCGCCGCTCTCCGCGCTCGCC 28932

RESULT 6

US-08-457-342-6
 ; Sequence 6, Application US/08457342
 ; Patent No. 5662898
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James M.
 ; APPLICANT: Beck, James Joseph
 ; APPLICANT: Hill, Dwight Steven
 ; APPLICANT: Ryals, John Andrew
 ; APPLICANT: Gaffney, Thomas Deane
 ; APPLICANT: Lam, Stephen Ting
 ; APPLICANT: Hammer, Phillip E.
 ; APPLICANT: Uknes, Scott Joseph
 ; TITLE OF INVENTION: Genes for the synthesis of
 ; TITLE OF INVENTION: antipathogenic substances
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Ciba-Geigy Corporation
 ; STREET: 7 Skyline Drive
 ; CITY: Hawthorne
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10532
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/457,342
 ; FILING DATE: 01-JUN-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/457,205
 ; FILING DATE: 01-JUN-1995
 ; APPLICATION NUMBER: 08/258,261
 ; FILING DATE: 08-Jun-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Elmer, James Scott
 ; REGISTRATION NUMBER: 36,129
 ; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919-541-8614
 ; TELEFAX: 919-541-8689
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 28958 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO

; ANTI-SENSE: NO
 ; US-08-457-342-6
 Alignment Scores:
 Pred. No.: 0.006 Length: 28958
 Score: 150.50 Matches: 255
 Percent Similarity: 34.66% Conservative: 111
 Best Local Similarity: 24.15% Mismatches: 405
 Query Match: 2.87% Indels: 292
 DB: 1 Gaps: 57
 US-09-929-769-7 (1-1029) x US-08-457-342-6 (1-28958)
 QY 14 ThrLeuGlyProProArgAlaAspSerGluPheGlnAlaLeuAsp----- 30
 Db 26222 ACCTCCGGCAACCGCTCTGTTCTCGAGCGCGAGACCGGCTCTCTGACGATGGGAT 26280
 QY 31 IleTrpPheProGluGluValProLeuPro-----ThrAlaPheValAspThrSer 48
 Db 26281 CGCTTCTTCGTGAGGTACGAGCCCGCATCCGCTCAAGCTCGGCTCCGAGACCTGC 26340
 QY 49 GluGluAlaLeuLeuProAspThrLeu-----LysLeuArgMetIleArgSerGluVal 67
 Db 26341 GAGGCTCACCGCTCATCCGCTCGCTCGCTCCATTCAGCGAGAGCCACCTC 26400
 QY 68 LeuArgLeuValAspAlaAlaLeuGlnAspLeuGluProGlnGlnLeuLeuPheVal 87
 Db 26401 GCCCGCTGCTC-----CTCTCCGCGGAGAGCTCTACCGAGCGCTCGACCTG 26457
 QY 88 GlnSerPheGlyIleProValSerMetSerLysLeuGlnPheLeuAspGlnAla 107
 Db 26458 AACGCTCTTCGCGCGCTCCCGCAAGCTCTCCACCTAC----- 26508
 QY 108 ValAlaHisAspProGlnThrLeuGlnGlnAsnIleMetAspLysAsnTyrMetAlaHis 127
 Db 26509 -----CCCTTCAACGAGGCTTCTGCTGAGAGCTTC-----ACGCGCTC 26553
 QY 128 LeuValGluValGlnHisGluArgGlyAlaSerGlyGlnThrPhe-----HisSer 145
 Db 26554 GCTGCGGAGCTC-----GCTCCGAGAGCTGAGCTCGCGGAGACCCG 26598
 QY 146 LeuLeuThrAlaSerLeuProProArgArgAsp----- 156
 Db 26599 CTGCTCGGCGCGCGCTGCG-----CTCGCGAGCGGATGCTTGTCTTCACAGAGCGCT 26657
 QY 157 -----SerThr 158
 Db 26658 CTCCTCGCAGAGCACCCGCTGCTCGAAGACCACTGCTTCGACATACCTGTCTGC 26717
 QY 159 GluAlaProLysProLysSerSerProGlu-----GlnProIleGlyGlnArgIleArg 177
 Db 26718 CAGGCGCGCGCTCGTGCAGTGTGCGCTCATGTGCGCCATCTCGCGCTCGACACGCT 26777
 QY 178 ValGlyThrGlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGln 197
 Db 26778 CGAAGACGTCAAGCTCGA----- 26795
 QY 198 IlePheProLeuSerProAspProArgTrpGlnSerSerSerProArgProValAlaLeu 217
 Db 26796 -----CCCCCCTGCTCTCTCCATCCGAGGCGCGCTCTCTCA--- 26837
 QY 218 AlaLeuGlnAlaLeuGlyGlnGluLeuAlaArgValValGlnGly-----SerProGlu 236
 Db 26838 GATCCGCTCGGCGCGCGGAGCGGAGG-----GAAAGCGGCTCTCCCTC 26887
 QY 237 ValProGlyIleThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHis 256
 Db 26888 ATACCGCGCGCGAGACCGCTTC-----AGATAGCCCTCGAGCTCGACGCCA--- 26938
 QY 257 GlyGlyAlaLeuValMetSerMetHisArgSerHisPheLeuAlaCys----- 272
 Db 26939 ---GGGCTCTCTCGCGAGCTAGACCGCTCC-CATTCTTCATGATGCTCCGCAATGCG 26994

QY 273 ---ProLeuLeuAArgLInLeuCySGInTyRGInAArgCyValProGInAspThr---Gly 290
 26995 CCCCC-----TCGGGCGCCACCAGGTGAGCCCAAGGT 27030
 QY 291 PheSerSerLeuPheLeuLysValLeuLeuGInMetLeuGInTrpLeuAspSerProGly 310
 27031 TTCACGCGACC-----CTCGAGACGCTGGG 27057
 QY 311 ValGluGlyPro-----LeuAArgAlaGInLeuAArgMet 322
 27058 CTTGCTTATGAGCCCGGAGTTCCAGGGCTTCGGCGCCGCTACAAACGGGCGACAGAGCT 27117
 QY 323 LeuAlaSerGlnAlaSerAlaGlyArg-----LeuSerAsp 335
 27118 CTTCCGCCGAGCCACAGCTCCGGACCGCGCGAAGAGACCGCGCTTTCCCTCCCA 27177
 QY 336 ValAArgGlyGlyLeuLeuAArgLeuAlaGlnAlaLeuAlaPheArg----- 350
 27178 CCCCCCGCTGCTGCACAGCGCTTCGAGGCGCTCGCTTTGAGACGACGACCAAGG 27237
 QY 351 ---GlnAspLeuGlnValAlaSerSerThrValAArgAlaValAlaLeuAlaThrLeuAArgSer 369
 27238 CTTCAAGATGCCCTTCTCGTGAGCGAGT-----ATCGCTGCGCTCCGCTCG 27285
 QY 370 GlyGluGlnCysSerValGluProAsp-----LeuIleSerLysValLeuGlnGly 386
 27286 -----GAGCCACCACCGCTGGCGCTGGCTTCCACCGCTCGAGGGCG 27327
 QY 387 LeuIleGlnValAArgSerProHisLeuGlnGlnLeuLeuThrAlaPhePheSerAlaThr 406
 27328 AATCCTCGGCGCTCGCTCC-----TCCTCGCGG 27354
 QY 407 AlaAspAlaAlaSerProPheProAlaCysLys-----ProValValAlaValSerSer 424
 27355 ACGCCAGAGCGCACCCATCGCTCGTCAAGCGCTCCGACATGCGCG----- 27402
 QY 425 LeuLeuLeuGlnGlnGlnLeuProLeuAlaGlyGlyLysProGlnAlaAspGlyLys 444
 27403 -----CCGCGTCCCGCGAGACCTCGCGAGACCCGGAGCG-----TCGCAC 27444
 QY 445 LeuGlnAlaValAArgLeuGlyProSerSerGlyLeuLeuValAspTrpLeuGlnMetLeu 464
 27445 CTTGATGCGCTC-----TTCCGATCGACTGAGCGAGCTGCA- 27482
 QY 465 AspProGlnValValSer-----SerCysProAspLeuGlnLeuAArgLeuLeu 480
 27483 AAGCCCACTACCGCCCATCGCGCCGAGCGGTGCTCTCGGCACAGAG----- 27536
 QY 481 PheSerAArgLysGlyGlnAlaGlnValProSerPheAArgProTrpLeuLeu 500
 27537 ---TCTCGACCTCGGACAGAGGT---GCCTCGACCGCTATACGACCTTCCTCTCT 27590
 QY 501 Thr-----LeuPheThrHisGlnSerSerTrpProThrLeuHisGln 514
 27591 ACGGAGCGCCCTCGACACAGGCGCTCGCTCCAAACCTCGTCATCGCCCTCATGCC 27650
 QY 515 CysIleAArgValLeuLeuGlyLysSerAArgGlnGln-----ArgPheAspPro 530
 27651 TCTGCCGAGGCGACCTCATCGCGAGCGCGCGGAGACACCGCGCGCTCGCTCG 27710
 QY 531 SerAlaSerLeuAspPheLeuTrpAlaCysIleHisValProAArgTrpGlnGlyAArg 550
 27711 CTTGCAAGCTG-----GCTCGCGACGAGCGCTCGCTCGCTCGCGCTCGC 27758
 QY 551 AspGlnAArgThrProGlnLysAArgAArgGlnGlnLeuValLeuAArgValGlnGlyProGln 570
 27759 ---CCTCGTACCGCGAGCGGCTCGCGCACCGCTGAAGAAGAGCTCAAGGCGCTCG- 27814
 QY 571 LeuIleSerLeuValGlnLeuLeuLeuAlaGlnAlaGlnThrAArgSerLysAspGlyAsp 590
 27815 CTCACGCGCGCTCTCGGCTCGCTCGCTCGCGCGAGAGAGACCCAGAGCGCGCTCG 27874
 QY 591 ThrAlaAlaCysSerLeuIleGlnAlaAArgLeuProLeuLeuLeuSerCysCysGly 610

Db 27875 TCGTCTCTCGACCTCGACGACGAGGAGGCTCCAGACGCGCTGTCGGCGCGCTCG 27934
 QY 611 AspAspGlnSerValAArgLysValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGly 630
 27935 ACGCAAGAGACGACGAGATCGCTCCGACAGGCAACCCCTCGTCCAA-----GGC 27988
 QY 631 AspSerValLeuGlyAArgAArgAspLeuLeuLeuGlnLeuTrpLeuGlnAArgPro 650
 27989 TCTCAGCGCTGCGCCGAGCGGCCACAG-----ACACAGCGCTGCC 28027
 QY 651 GlnLeuAArgValProValProGlnValLeuLeuHisSerGlnGlyAlaAlaSerSerSer 670
 28028 CCGAGGCGCTCGAGGAGCGAGCTGTCATACGGGAGGAGGACCGGCGCTCGCGCGCTCG 28086
 QY 671 ValCysLysLeuAspLysLeuIleHisAArgPheIleThrLeuLeu----- 685
 28087 GTCCGCGCGCGCTCGCTCGTAAACGACGCGCAACGACCTGCTGCTCACTCGCGCGCG 28146
 QY 686 -----AlaAspThrSerAspSerAArgAlaLeuGlnAsnAArgGlyAla 699
 28147 GCGGCGAGCGCTCGCGGCTGCTGATGCTTGCAGAGCGAG-----CTGAACTCTGGGCGCT 28203
 QY 700 AspAlaSerMetAlaCysAArgLysLeuAlaValAlaHisProLeuLeu-----LeuAArg 718
 28204 TCGGTACCGCTCGCGCTCGCGACGCGGCGATCCAGCGCGCTTAAGAGACTTCTGAT 28263
 QY 719 HisLeuProMet-----IleAlaAlaLeuLeuHisGlyAArgThrHisLeuAsnPhe 735
 28264 AACATTCGACGCGCTCACCGCGCTCGCGCGCGCGCTGTCATGCGCGCGCGCTCGAGCG 28323
 QY 736 GlnGluPheAArgGlnGlnHisHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlu 755
 28324 GAT-----CTGCTCGCGCGCCATGAGC 28344
 QY 756 LeuLeuGln-----ProHisValPheAArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeu 774
 28345 CTCGAGCGGATGACCGCGCTTCTCGCCGCCCAAGATGATGCCGCG-----TGGCAC 28395
 QY 775 LeuSerPheIleAArgLeuLeuAsnTyAArgLysSerSerAArgHisLeuAlaAlaPhe 794
 28396 -----TTGCATCAGCTCACCCAGATGAAGCGCTCGCGCTCG 28434
 QY 795 Ile-----AsnLysPheValGlnPheIleHisLysTyIleThrTyTrsnAlaProAla 812
 28435 ATCTCTTCTGCTCGCTCGCGCGCGCTCGCGAGCTCGAGCTCACTCACTACAGCGCTCG 28494
 QY 813 AlaIleSerPheLeuGlnLysHisAlaAsp-----ProLeuHis 825
 28495 GCGAGCGCTTCTCTCGATGCGCTGCGCACACCGCGCGCGAGGCGCTCGCTCGCTCA 28554
 QY 826 AspLeuSerPheAspAsn-----SerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeu 844
 28555 TCGCTCGCGCTGAGCAGCTCGGCGCGCGCGAGCGCAATGACAGAGACCTGACGCGCGCG 28614
 QY 845 SerLeuProSerAArgAspAArgTrpAArg-----GlyLeuAspGlnGlnGlyGlnGlu 863
 28615 GCGCGCGCT-----CGCATGAGCGCGCGCGCGCTTCCCTCGACCTCTGAGGAG 28662
 QY 864 GlnSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThr 882
 28663 AGG-----CTCGCGCTCTCGATGCGCGCGCTTCTCGGACCGAGCGCGCG 28707
 QY 883 -----AlaAlaGlnMetAlaPro 888
 28708 CTTGTCGCCGCGCGCTTCACTTGAACGCGCTCAGGCGGAGAGCGGCGAGCTGCCCGG 28767
 QY 889 TyrMetLysAArgLeuSerAArgGlyGlnThrVal----- 899
 28768 TTGTTCAACAGTCTGCTCGCTCGCTCGACCGCTGACGAGCGCGCGAGCGCGCGCGAG 28827
 QY 900 ---GlnAspLeuGlnGlnValLeuSerAspIleAspGlnMetSerAArgAArgProGln 918

```

Db 28828 GCCTGCGTCTTACAGAGGGCCTCTCAGGCCCTCCGGCCCGCCGAGACGGAGCGT---GCC 28884
Oy 919 IleuSerPhepheSerThrAsnLeuGlnArgLeuMetSerSeraLa 934
      :::::
Db 28885 CTGCTCGATCTCATTCACGACCGAGCCGCCGCTGCTCGGCGCTCGCC 28932
      :::::

RESULT 7
US-08-457-646A-6
: Sequence 6, Application US/08457646A
: Patent No. 5679560
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Ligott, James M.
: APPLICANT: Beck, James Joseph
: APPLICANT: Hill, Dwight Steven
: APPLICANT: Ryals, John Andrew
: APPLICANT: Gaffney, Thomas Deane
: APPLICANT: Lam, Stephen Ting
: APPLICANT: Hammer, Phillip E.
: APPLICANT: Uknes, Scott Joseph
: TITLE OF INVENTION: Genes for the synthesis of
: TITLE OF INVENTION: antipathogenic substances
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cliba-Geigy Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: NY
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/457,646A
: FILING DATE: 01-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/457,205
: FILING DATE: 01-JUN-1995
: APPLICATION NUMBER: 08/258,261
: FILING DATE: 08-JUN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Elmer, James Scott
: REGISTRATION NUMBER: 36,129
: REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8614
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 28958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
US-08-457-646A-6

Alignment Scores:
Pred. No.: 0.006 Length: 28958
Score: 150.50 Matches: 255
Percent Similarity: 34.66% Conservative: 111
Best Local Similarity: 24.15% Mismatches: 405
Query Match: 2.87% Indels: 292
DB: 1 Gaps: 57
US-09-929-769-7 (1-1029) x US-08-457-646A-6 (1-28958)
Oy 14 ThrLeuGlyProProArgAlaAspAspSerGluPheGlnAlaLeuLeuAsp----- 30

```

D 27118 CTTCCGCGAGCGCAAGCTCCCGAGCGCGCGAGAGAGAGCGCGCTTTGCCCTCCA 27177
OY 336 ValArgGlyGlyLeuLeuArgLeuAlaLeuAlaPheArg----- 350
D 27178 CCCCCTGCTGCGACAGCGCTTTCAGAGCGCTCGCTTTGAGAGCGAGCAAGCG 27237
OY 351 ---GlnAspLeuGluValValSerSerThrValArgAlaValIleAlaThrLeuArgSer 369
D 27238 CTTGAGAGATGCTCTGCTGAGCGAGT-----ATGCGTGGCTCCGCTCG 27285
OY 370 GlyGlyGlnGlySerValGluProAsp-----LeuIleSerIysValLeuGlnGly 386
D 27286 -----GAGCCACCGCCCTGCGCTGCTTCCACCGCTCTGAGGCGC 27327
OY 387 LeuIleGluValArgSerProHisLeuGlnGluLeuLeuThrAlaPhePheSerAlaThr 406
D 27328 ATCTCTCGGCTGCTC-----TCCTCGCGC 27354
OY 407 AlaAspAlaIleSerProPheProAlaCysLys-----ProValValValSerSer 424
D 27355 ACGCGAGAGCGAGACCATCGCTCGTCAAGCGCTCGCATGCGCG----- 27402
OY 425 LeuLeuLeuGlnGluGluProLeuAlaGlyGlyLysProGlyAlaAspGlySer 444
D 27403 -----CCGCGTCCGCGAGCACTCCGAGACCGCGAGCG-----TCCAC 27444
OY 445 LeuGluAlaValArgLeuGlyProSerSerGlyLeuLeuValAspThrLeuGluMetLeu 464
D 27445 CTCGATGCGCTC-----TTCCGATGCGATGAGCGAGCGTGCAT 27482
OY 465 AspProGluValValSer-----SerCysProAspLeuGlnLeuArgLeuLeu 480
D 27483 AGCCCGACCTCAGCGCCATCGCCCGAGCGCGCTCTCCGAGAGAG----- 27536
OY 481 PheSerArgArgGlyGlyGlnAlaGlnValProSerPheProGlyLeuLeu 500
D 27537 ---TCTCGACCTCGGAGCAGGCT---GCTCTCGACCGCTATACGAGCTTCTCTCT 27590
OY 501 Thr-----LeuPheThrHisGlnSerSerThrProThrLeuHisGln 514
D 27591 ACGAGCGCGCTCAGACAGCGCGCTTCCAGCTCGATCGCGCTTCATCGC 27650
OY 515 CysIleArgValLeuLeuGlyLysSerArgGluGln-----ArgPheAspPro 550
D 27651 TCTGCCGAGGCGACCTCATCGAGCGCGCGAGACCGCGAGCGCTCGCTCT 27710
OY 531 SerAlaSerLeuAspPheLeuThrAlaCysIleHisValProArgIleThrGlnGlyArg 550
D 27711 CTTGCAAGCTG-----GCTCGCGAGAGCGCGCTCTCGCGCGCTCGC 27758
OY 551 AspGlnArgThrProGlnLysArgArgGluGluLeuValLeuArgValGlnGlyProGlu 570
D 27759 ---CCTCGCTCAGCGCGAGCGCTCGCGACCGCTCAAGAGAGCTGAGAGCGCTCG- 27814
OY 571 LeuIleSerLeuValGluLeuIleLeuAlaGluThrArgSerGlnAspArgValAsp 590
D 27815 CTCAGCGCGCTCTGCGGCTCTCGCTCGCTCGCGAGAGAGCGAGCGCGCTCT 27874
OY 591 ThrAlaAlaCysSerLeuLeuGlnAlaArgLeuProLeuLeuLeuSerCysCysGly 610
D 27875 TCGTCTCTCGACCTCGAGAGAGAGCGAGCGCTCCAGAGCGCGCTCTCGCGCTCG 27934
OY 611 AspAspGluSerValArgLysValThrGlnHisLeuSerGlyCysIleGlnGlnArgGly 650
D 27935 ACGCAAGAGAGCGAGATCGCTCGCGACCGCAAAACCGCTCTGTTCAA-----GGC 27988
OY 631 AspSerValLeuGluArgArgCysArgAspLeuLeuGlnLeuLeuLeuGlnArgPro 650
D 27989 TCTCAGCGCTCGCGAGGCGCGCGAG-----ACACGCGCTCGC 28027
OY 651 GluLeuArgValProValProGluValLeuLeuHisSerGluValAlaAlaSerSer 670
D 28028 CCGAGAGCGCTCGAGAGCGAGCTGCTCATTCAGCGAGAGAGCGAGCGCGCTCGCGCGCTG 28086

OY 671 ValCysLysLeuAspGlyLeuLeuIleIleArgPheIleThrLeuLeu----- 685
D 28087 GTGCGCGCGCGCTGCTGCTTAACCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 28146
OY 686 -----AlaAspThrSerAspSerArgAlaLeuGlnAsnArgGlyAla 699
D 28147 GCGCGAGCGCTCGCGGCTGTATGCTTCTGCGAGAGAG---CTCGAGCTCTGGGCGCT 28203
OY 700 AspAlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeu---LeuLeuArg 718
D 28204 TCGGTACCGCTGCGCGCGCTGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 28263
OY 719 HisLeuProMet-----IleAlaAlaLeuLeuHisGlyArgThrHisLeuAsnPhe 735
D 28264 AACATTCCAGCGCTCAGCGCGCGAGCGCGCGCTGCTCATGCGCGAGCGAGCGAGCG 28323
OY 736 GlnGluPheArgGlnGlnHisHisLeuSerCysPheLeuHisValLeuGlyLeuGly 755
D 28324 GAT-----CTGCTCGCGCGCGATGAGC 28344
OY 756 LeuLeuGln---ProHisValPheArgSerGlnHisGlnGlyAlaLeuThrAspCysLeu 774
D 28345 CTCGAGCGAGTACGCGCGCTTCTCGCGCGCGCGAGATGAGCGCG---TGGCAC----- 28395
OY 775 LeuSerPheIleArgLeuLeuLeuAsnArgLysSerSerArgHisLeuAlaAlaPhe 794
D 28396 -----TTCATCAGCTCAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 28434
OY 795 Ile-----AsnLysPheValGlnPheIleHisLysTyIleThrTyrrAsnAlaProAla 812
D 28435 ATCTCTTCTCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 28494
OY 813 AlaIleSerPheLeuGlnLysHisAlaAsp-----ProLeuHis 825
D 28495 GCGAGCGCGCTTCTGATGCGCTTTCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 28554
OY 826 AspLeuSerPheAspAsn---SerAspLeuValMetLeuLysSerLeuAlaGlyLeu 844
D 28555 TCCCTCGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 28614
OY 845 SerLeuProSerArgAspArgThrAspArg---GlyLeuAspGluGlnGlyGlnGly 863
D 28615 GCGCGCGCT-----GCGATGAGAGCGCGCGCGCTTCTCGAGCTGAGAG 28662
OY 864 GluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThr--- 882
D 28663 AGG-----CTCGCGCTTCTGATGCGCGCGCTTCTCGAGCGAGAGCGCGCG 28707
OY 883 -----AlaAlaGluMetAlaPro 888
D 28708 CTGCTCCGCGCGCTTCTGAGCTTTCAGCGCGCTGAGCGCGAGCGAGCGAGCGAGCG 28767
OY 889 TyrMetLysArgLeuSerArgGlyGlnThrVal----- 899
D 28768 TTGTTCCAGCTGCTGCTGCGCGCTGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 28827
OY 900 ---GlnAspLeuGluValLeuLeuSerAspIleAspGluMetSerArgArgArgProGlu 918
D 28828 GCTCTGCTTACAGAGCGCGCTTCTGAGCGCGCTTCTCGAGCGCGAGCGAGCGAGCG 28884
OY 919 IleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAla 934
D 28885 CTGCTGATCTCATCGCGAGCGAGCGCGCGCTGCTGCGCGCTGCGC 28932
RESULT 8
US-08-458-076A-6
Sequence 6, Application US/08458076A
Patent No. 5698425
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph


```

APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
TITLE OF INVENTION: antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESS: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,076A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REFERENCE/DOCKET NUMBER: 36,129
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
FAX: 919-541-8614
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-458-076A-6

Alignment Scores:
Pred. No.: 0.006 Length: 28958
Score: 150.50 Matches: 255
Percent Similarity: 34.66% Conservative: 111
Best Local Similarity: 24.15% Mismatches: 405
Query Match: 2.87% Indels: 292
Gaps: 57

US-09-929-769-7 (1-1029) x US-08-458-076A-6 (1-28958)
Qy 14 ThrLeuGIYProPArgAlaAspSerGIuPheGlnAlaLeuLeuAsp----- 30
Db 26222 ACCCTCGGCAACCGTCTCTGTTTC-TGAGAGCGGACCGGCGCTCGACATGGGCAT 26280
Qy 31 lIetPheProGIuGIuLysProLeuPro-----ThrAlaPheLeuValAspThrSer 48
Db 26281 CGCTCTCTGTCGAGGTGACGCCCCCATCCGTCACAGCTGCGCTCCGCGAGACCTGC 26340
Qy 49 GIuGIuAlaLeuLeuLeuProAspTrpLeu---LysLeuArgMetIleArgSerGIuVal 67
Db 26341 GAGCGCTCGACGTCGATCCGCTCGCTCGCTCATTCGACGCGAGCAAGCCACCTGC 26400
Qy 68 LeuArgLeuValAlaAlaLeuLnsPleuGIuProGlnGlnLeuLeuLeuPheVal 87
Db 26401 GCCCGCTGCTC---CTCTCTGGGCGGAGCTCTTACCCGAGGCGCTGCGCTGACTGC 26457

```

```

Qy 88 GlnSerPheGlyIleProValSerSerMetSerLysLeuLeuGlnPheLeuAspGlnAla 107
Db 26458 AACGCGCTTCTGCGGCGCCCTGCTCCCGCAAGGTCCTCCCTCCACACCTAC----- 26508
Qy 108 ValAlaHisAspProGlnThrLeuGIuGlnAsnIleMetAspLysAsnTrpMetAlaHis 127
Db 26509 -----CCCTTCACACGCGAGCGCTTCTGCTGACGAGCCCTCC---ACGCGGAC 26553
Qy 128 LeuValGIuValGlnHisGlnArgGlyAlaSerGlyGIuThrPhe-----HisSer 145
Db 26554 GCTCGCGAGCTC-----GCTTCGCGAGGCTGACCTCGGCGCACACCCG 26598
Qy 146 LeuLeuThrAlaSerLeuProProArgArgAsp-----SerThr 156
Db 26599 CTGCTCGGCGGCGCGCGCTGC-CCTGCGCGACCGCGATGCTTGTCTTCACAGACGGCT 26657
Qy 157 -----SerThr 158
Db 26658 CTCCCTCGCAGACACCGCGTGCAGAGACACGCTCTTCGCGATACCTGCTCTGC 26717
Qy 159 GIuAlaProLysProLysSerSerProGIu---GlnProIleGlyGlnGlyArgIleArg 177
Db 26718 CAGGCGCGCTCTCGAGCTCGCCCTGCATGCTGCGCTGCGCTGCGACACGCT 26777
Qy 178 ValGIYThrGlnLeuArgValLeuGIYProGIuAspAspLeuAlaGIYMetPheLeuGln 197
Db 26778 CCAAGACGTCACGCTCGA----- 26795
Qy 198 lIePheProLeuSerProAspProArgTrpGlnSerSerProArgProValAlaLeu 217
Db 26796 -----CCCCCTCGCTCTCCATCGAGCGCGCTCTCTCA--- 26837
Qy 218 AlalaGlnGlnAlalaLeuGIYGlnGluLeuAlaArgValAlaGlnGly---SerProGIu 236
Db 26838 GATCTCGCGGCGCGCGCGAGGCTG-CGTGAC-----GAAGGCGCTCTCTCGTTC 26887
Qy 237 ValProGIYlIeThrValArgValLeuGlnAlaAlaThrLeuLeuSerSerProHis 256
Db 26888 ATAGCGCGCGCGCGCGCGCTTC-----AGATGCGCGCGCGCTCGCGCGCA--- 26938
Qy 257 GIYGIYAlaLeuValMetSerMetHisArgSerHisPheLeuAlaGly----- 272
Db 26939 ---GCGGCTCTCTCGCGAGCTAGCCCTGC-CATTCGCTCGATGCTCCCGCATATGCG 26994
Qy 273 ---ProLeuLeuArgGlnLeuGlyGlnTrpGlnArgCysValProGlnAspThr---Gly 290
Db 26995 CCCCC-----TCGGCGCGCACCCAGGTGACACCAAGGT 27030
Qy 291 PheSerSerLeuPheLeuLysValLeuLeuGlnMetLeuGlnTrpLeuAspSerProGIY 310
Db 27031 TTCTAGCGAGCC-----CTCAGAGCGCGTGG 27057
Qy 311 ValGIuGIYGIYPro-----LeuArgAlaGlnLeuArgMet 322
Db 27058 CTTGCTTATGAGCCCGAGTTCAGGCGCTCCGCCGCTCTACAGCGCGGAGAGCT 27117
Qy 323 LeuAlaSerGlnAlaSerAlaGlyArgArg-----LeuSerAsp 335
Db 27118 CTTGCGCGAGCAAGCTCCGCGAGCGCGCGCAAGAGAGCGCGCTGTTTGCCCTCA 27177
Qy 336 ValArgGIYGIYLeuLeuArgLeuAlaGlnAlaLeuAlaPheArg----- 350
Db 27178 CCGCGCGCTGTCGAGAGCGCTTCGAGCGCTGCTGTGTAAGACGACGCAAGCAAGGC 27237
Qy 351 ---GlnAspLeuGlnValValSerSerThrValArgAlaValIleAlaThrLeuArgSer 369
Db 27238 CTTGAGAGTGCCTCTCTGAGCGGAGT-----ATGCTGCGCTCGCGCTG 27285
Qy 370 GIYGIuGlnCysSerValGIuProAsp-----LeuIleSerLysValLeuGlnGly 386
Db 27286 -----GAGCGACACCGCTCGCGGTGCTTTCACGCTCTGAGAGGCG 27327

```

```

QY 387 LeuIIleuIValArgSerProHISleuGIuLeuLeuThraIaPheSerAlaThr 406
      |||||
Db 27328 MATCTCGCGCTCGCTC-----TCCTCGCGC 27354
QY 407 AlaAspAlaIaSerProPheProAlaCysIys-----ProValValValSerSer 424
      |||||
Db 27355 ACCGCGAGGCGGAGGCGGCGCTCGGCGAGCGCTCGGCGAGCGCGG 27402
QY 425 LeuLeuLeuGIuGIuGIuProLeuAlaGIuGIuLysProGIuAlaAspGIuLys 444
      |||||
Db 27403 -----CCGCGCGCGCGAGAGCTCGGCGAGCGCGGAGG-----TCCAC 27444
QY 445 LeuGIuAlaValArgLeuGIuProSerSerGIuLeuValAspTrpLeuGIuMetLeu 464
      |||||
Db 27445 CTGATGCGCTC-----TTCGCGATCGACTGCGAGCGCTCA- 27482
QY 465 AspProGIuValValSer-----SerCysProAspLeuGIuLeuArgLeu 480
      |||||
Db 27483 AGCGCCACCTCAGCGCGCGCGCGCGCGGCTCGCTCGCGCAGACAG----- 27536
QY 481 PheSerArgArgLysGIuLysGIuGIuAlaGIuValProSerPheArgProTrpLeu 500
      |||||
Db 27537 ---TCTGACCTCGGAGCGAGGT---GCCTCGACCGCTATACGACCTTGCTGCT 27590
QY 501 Thr-----LeuPheThrHISGIuSerTrpProThrLeuHISGIu 514
      |||||
Db 27591 AGCGAGCGCGCTCGACGAGGCGCTCGCTCGCTCGACGCTCGATCGCGCTTCATCGC 27650
QY 515 CysIleArgValLeuGIuGIuLysSerArgGIuGIu-----ArgPheAspPro 530
      |||||
Db 27651 TCGCGCGGAGGCGAGCTCTCGCGAGCGCGCGCGCGAGACCGCGCGCGCTCGCT 27710
QY 531 SerAlaSerLeuAspPheLeuThrAlaCysIleHISValProArgIleTrpGIuArg 550
      |||||
Db 27711 CTGCAAGCGCTG-----GCTCGCGGAGGAGCGCGCTCGCTCGCGCGCTCGC 27758
QY 551 AspGIuArgThrProGIuLysArgArgGIuGIuLeuValLeuArgValGIuGIuProGIu 570
      |||||
Db 27759 ---CTGCTGACCGCGAGCGCGCTCGCGCGCGCGCGCGCGAGCGCGCGCGCTCG- 27814
QY 571 IleIleSerLeuValGIuLeuIleLeuAlaGIuAlaGIuThrArgSerGIuAspGIu 590
      |||||
Db 27815 CTGACGCGCGCTCTCGGCGCTCGCTCGCTCGCGCGCGAGACGAGACCGCGCGCTCG 27874
QY 591 ThrAlaAlaCysSerLeuIleGIuAlaArgLeuProLeuLeuSerCysCysGIu 610
      |||||
Db 27875 TCGCTCGCTGCGACCTCGACGAGCGAGCGCGCTCGCGAGCGCGCTCGCGCGCTCG 27934
QY 611 AspArgGIuSerValArgValThrGIuHISLeuSerGIuLysIleGIuGIuTrpGIu 630
      |||||
Db 27935 AGCGAGAGGCGCGAGAGTGCCTCGCGAGCGCGCGCGCGCGCGCGCTCGCTCGCA- 27988
QY 631 AspSerValLeuGIuArgArgCysArgAspLeuLeuLeuGIuLeuGIuArgPro 650
      |||||
Db 27989 TCTGACGCGCTCGCGCGCGCGCGCGCGCG-----ACACGCGCGCTCG 28027
QY 651 GIuLeuArgValProValProGIuValLeuLeuHISSerGIuAlaAlaIaSerSer 670
      |||||
Db 28028 CCGCAGCGCGCTCGGAGCGCGCGCTCGCATACGCGAGCGAGCGCGCGCGCGCTCG 28086
QY 671 ValCysLysLeuAspLeuLeuIleHISArgPheIleThrLeu----- 685
      |||||
Db 28087 GTTCGCGCGCGCGCTCGTGAACGAGCGAGCGCGCGCGCGCGCGCTCGCTCGCGCGCG 28146
QY 686 -----AlaAspThrSerAspSerArgAlaLeuGIuGIuAsnArgGIuAla 699
      |||||
Db 28147 GCGCGCAGCGCGCTCGGCTGTGTCTCTCGGAGCGAG-----CTCGAGCTCTCGGCGCT 28203
QY 700 AspAlaSerMetAlaCysArgGIuLysLeuAlaValAlaHISProLeuLeu---LeuLeuArg 718
      |||||
Db 28204 TCGGTCACCGCTCGCGCGCGCGAGCGAGCGCGCGCGCTCGAAGAGCGCTCGGAGT 28263
QY 719 HISLeuProMet-----IleAlaAlaLeuLeuHISGIuArgThrHISLeuAsnPhe 735

```

```

Db 28264 AACATTCGAGCGCGCTCGACCGCGCGCGCTCGGATCGCGCGCGCGCTCGAGCGC 28323
QY 736 GIuGIuPheArgGIuGIuAsnHISLeuSerCysPheLeuHISValGIuLeuGIu 755
      |||||
Db 28324 GAT-----CTGCTCGCGCGCGATGAGC 28344
QY 756 LeuLeuGIu---ProHISValPheArgSerGIuHISGIuAlaLeuTrpAspCysLeu 774
      |||||
Db 28345 CTGAGCGGAGTACCGCGCGCTCGCGCGCGCGCGCGAGTACGATCGCGCG---TGGCAG- 28395
QY 775 LeuSerPheIleArgLeuLeuLeuAsnTrpArgLysSerArgArgHISLeuAlaIaPhe 794
      |||||
Db 28396 -----TTCATACAGCTACCGCGAGATAGCGCTTGCGCGCTTC 28434
QY 795 Ile-----AsnLysPheValGIuPheIleHISLysTrpIleThrTrpAsnAlaProAla 812
      |||||
Db 28435 ATCTCTTCTCGTCGCGCGCGCGCGCGCTCGGACGCTACGATCGAATCGAGCGCT 28494
QY 813 AlaIleSerPheLeuGIuLysHISAlaAsp-----ProLeuHIS 825
      |||||
Db 28495 GCGAGCGCGCTCTCGATCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCTCGCTCA 28554
QY 826 AspLeuSerPheAsn---SerAspLeuValMetLeuLysSerLeuLeuAlaGIuLeu 844
      |||||
Db 28555 TCGCTCGCGTGAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 28614
QY 845 SerLeuProSerArgAspAspArgThrAspArg---GIuLeuAspGIuGIuGIuGIu 863
      |||||
Db 28615 GCGCGCGCT-----CGCATGAGCGCGCGCGCGCTCGCTCGCGCGCGCGCGCGAG 28662
QY 864 GluSerSerAlaGIuSerLeuProLeuValSerValSerLeuPheThrProLeuThr--- 882
      |||||
Db 28663 AGG-----CTCGCGCTCTCGATGCGCGCGCTCTCGCGAGCGAGCGCGCG 28707
QY 883 -----AlaAlaGIuMetAlaPro 888
      |||||
Db 28708 CTGCTCGCGCGCGCGCTTCGACTTGAGCGCGCGCTCGAGCGCGCGCGCGCGCGCG 28767
QY 889 TyrMetLysArgLeuSerArgGIuGIuThrVal----- 899
      |||||
Db 28768 TTGTTCAGAGCTCGTCGCGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCG 28827
QY 900 ---GluAspLeuLeuGIuValLeuSerAspIleAspGIuMetSerArgValArgProGIu 918
      |||||
Db 28828 GCGCTCGCTGCTACAGCGCGCGCTCTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 28884
QY 919 IleLeuSerPhePheSerThrAsnLeuGIuArgLeuMetSerAla 934
      |||||
Db 28885 CTGCTGATGTCATCGCGAGCGAGCGCGCGCGCGCGCGCGCTCGCGCGCTCGCG 28932

```

RESULT 9
 US-08-764-233A-4
 Sequence 4, Application US/08764233A
 Patent No. 5716849
 GENERAL INFORMATION:
 APPLICANT: Ligon, James M.
 APPLICANT: Schupp, Thomas
 APPLICANT: Beck, James J.
 APPLICANT: Hill, Dwight S.
 APPLICANT: Neff, Snezana
 APPLICANT: Ryals, John A.
 TITLE OF INVENTION: Genes for The Biosynthesis of Soraphen
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 520 White Plains Road, P.O. Box 2005
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ. ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1
OS-08-764-233A-4

```

Alignment Scores:		
Pred. No.:	0.006	2895
Score:	150.50	Matches: 255
Percent Similarity:	34.66%	Conservative: 111
Best local Similarity:	24.15%	Mismatches: 405
Query Match:	2.87%	Indels: 292
DB:	1	Gaps: 57

OY	465	AspProGluValValSer-----SerCysProAspLeuGlnLeuArgLeuLeu	480
Db	27483	AAAGCCCACTACCGCCCATCGCCCGCAGCGCGTGCCTTCGGCAGCAAGAG-----	275386
OY	481	PheSerArgArgGlyGlyGlyGlnGlnGlnValProSerPheArgProTyrLeuLeu	500
Db	27537	---TCTCGACCTCGGAGACCAGAGT---GGCTCTCGACCGCATACCGACTTGGCTGCT	275906
OY	501	Thr-----LeuPheThrIleGlnSerSerTyrProThrLeuHisGln	514
Db	27581	ACGCAAGCGCCCTCGACCAAGCGCGTTCGCTCCAAAGCTCTGTATCGCCCTTCATCCG	276506
OY	515	CysIleArgValIleLeuGlnGlyLysSerArgGlnGln-----ArgPheAspPro	530
Db	27651	TCTGCCCAAGGCGACGACTTCATCGCAGAGCGCGCCGCGCAGACGACCGGCGCTCGCCT	277106
OY	531	SerIleSerLeuAspPheLeuTyrAlaCysIleHisValProArgIleTyrPheGlnArg	550
Db	27711	CTTGCAGACCTG-----GCTGCACGACAGCGCGCTCGGCTCTCGCGCTCGC	277586
OY	551	AspGlnArgThrProGlnLysArgArgGlnGlnLeuValLeuArgValGlnGlyProGln	570
Db	27759	---GCTGCTACCCGACGCGCGCTGCGCACCCGCTGAAAGAGAGTAAAGGCGCTCG-	278146
OY	571	LeuIleSerLeuValGlnLeuIleLeuIleGlnAlaGlnAlaGlnThrArgSerGlnAspLys	590
Db	27815	CTCAGCGCGCTCTGGGGTCTCGCTCGCTCGCGCAGACGACAGACCCAGAGCGCGCTC	278746
OY	591	ThrAlaIleArgSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCysGly	610
Db	27875	TCTGCTCTGCTGACTCTCGACGACGACGCGGCTCCAGCAGCGCTCTCGGCGCGCTCG	279346
OY	611	AspArgGluSerValArgGlyValThrGlnHisLeuSerGlyCysIleGlnGlnTyrGly	630
Db	27935	ACCGAAGAGCGGACGAAATCGCGCTCGGACGCGCAACCCCTCGTCCAA-----GGC	279886
OY	631	AspSerValLeuGlyArgArgCysArgAspLeuLeuGlnLeuTyrLeuGlnArgPro	650
Db	27989	TCTCAGCGCTGCCCGGCGCGCGCCACGCG-----ACACAGCGCTCCG	280276
OY	651	GluLeuArgValProValProGluValLeuLeuHisSerGlnGlyAlaAlaSerSerSer	670
Db	28028	CCGCAAGCGCTCGGAGGACCGCTC-CTCATCAGGAGGACGACCGGCGCTCGGCGCTCG	280866
OY	671	ValCysLysLeuAspGlyLeuIleHisArgPheIleThrLeuLeu-----	685
Db	28087	GTCGCGCGCGCGCTGCTGAACACGACGCGGCAAGCACTGCTCTCTACCTCGCGCGAC	281486
OY	686	-----AlaAspThrSerAspSerArgAlaLeuGlnAlaSerArgGlyAla	699
Db	28147	GAGCGGAGCGCTCGGCGTGTATGCTTCTTGGACGAG-CTGAGACCTCGGCGGCT	282036
OY	700	AspAlaSerMetAlaCysArgGlyLysLeuAlaValAlaHisProLeuLeu--LeuLeuArg	718
Db	28204	TCCGTCACCGCTCGCGCGCGAGCGAGCTGGCGGATTCGACGCGGCTCTTAAGGACCTTCTGAT	282636
OY	719	HisLeuProMet-----IleAlaIleLeuLeuHisGlyArgThrHisLeuAspPhe	735
Db	28264	AACATTTTGGAGCGCTACCGCGGTGCGCGCGCTGTCGACGCGCGGACGCTCTCGACGCG	283236
OY	736	GlnGluPheArgGlnGlnAlaHisIleLeuSerCysPheLeuHisValLeuGlyLeuLeuGln	755
Db	28324	GAT-----CTGCTCGGCGCCATGAC	283446
OY	756	LeuLeuGln--ProHisValPheArgSerGlnHisGlnGlyAlaLeuTyrAspCysLeu	774
Db	28345	CTCGAGCGGATCGACCGGCTTTCGCGCCCAAGATCATGCGCGC---TGGCAGC-----	283956
OY	775	LeuSerPheIleArgLeuLeuLeuSerGlnTyrArgLysSerSerArgHisLeuAlaPhe	794
Db	28396	-----TGGATCAGCTACCCCAAGATACCCCTTTCGCGCTTCG	284346

OY		795	Ile-----AsnLysPheValGlnIleProHisLysTyrLlleThrLysAlaProAla	812
Db		28435	ATCCTCTTCGTCCCGTGCGGGCGCTCCCGGACAGCTCAAGTGCACACTAACGCCGT	28496
OY		813	AlaIleSerPheLeuGlInylHisAlaSp-----ProLeuHis	825
Db		28495	GCGACCCTTCCTCGATCGAGCTTCCAGCACACCACCGCGCGCGCAAGGCGCTCCCTGCCTCA	28554
OY		826	AspLeuSerPheAlaSpApsr---SeraspLeuValMetLeuLysSerLeuLeuAlaGlyLeu	844
Db		28555	TCCCTGGCTGGAGCCACTGGCGCGCGGCGCCACCGCATACAGACAGCTCAGCGCCGC	28612
OY		845	SerLeuProSerArgAspaSpargThrasParq---GlyLeuAspLnuGluGlyGlu	863
Db		28615	GCGCGCCCT-----CGCATGGAGCGCGCGCGCTTCCTCCGACCTGTGGAGAG	28665
OY		864	GluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuthr---	882
Db		28663	AGG-----CTCCGCTCTTCGATGGGCGGCTTTCCGAACCGAGACCGCC	28707
OY		883	-----AlalaGluMetAlaPro	888
Db		28708	CTGGTCCCCGCGCGCTTCGACTTGAGCGCGCTCAGGCGGAGAAGCGCGAGCTCCCCCG	28766
OY		889	TyrmetylsArgLeuSerargylGlnThrval-----	899
Db		28768	TGTTCACAACGCTCTCGTCCGCGCTCGACACCTTAGCAGMAAGCGCGCCAGAACCGCCCAG	28827
OY		900	---GluasPheLeuGluValLeuSeraspIleAspGluMetSerArgArgProGlu	918
Db		28828	GCCCTGCGCTTACAGAGCGCTCTCAGCCCTCCGCGCGCGCGAGACGCAAGCT---GCC	28886
OY		919	IleLeuSerPhePheSerThrasInleuglNalArgLeuMetSerSerAla	934
Db		28885	CTGCTCGATCTCATCCGACCGAAGCGCGCGCTTCGCGCTCGCGCTGCC	28932
	RESULT 10			
	US-08-457-335A-6			
/	Sequence 6, Application US/08457335A			
/	Patent No. 5723759			
	GENERAL INFORMATION:			
	APPLICANT: Schupp, Thomas			
	APPLICANT: Ligon, James M.			
	APPLICANT: Beck, James Joseph			
	APPLICANT: Hill, Dwight Steven			
	APPLICANT: Ryals, John Andrew			
	APPLICANT: Gaffney, Thomas Deane			
	APPLICANT: Lam, Stephen Ting			
	APPLICANT: Hammer, Phillip E.			
	APPLICANT: Uknes, Scott Joseph			
	TITLE OF INVENTION: Genes for the synthesis of			
	TITLE OF INVENTION: antipathogenic substances			
	NUMBER OF SEQUENCES: 22			
	CORRESPONDENCE ADDRESS:			
	ADDRESSEE: Cliba-Gelgy Corporation			
	STREET: 7 Skyline drive			
	CITY: Hawthorne			
	STATE: NY			
	COUNTRY: USA			
	ZIP: 10532			
	COMPUTER READABLE FORM:			
	MEDIUM TYPE: floppy disk			
	COMPUTER: IBM PC compatible			
	OPERATING SYSTEM: PC-DOS/MS-DOS			
	SOFTWARE: PatentIn Release #1.0, Version #1.25			
	CURRENT APPLICATION DATA:			
	APPLICATION NUMBER: US/08/457, 335A			
	FILING DATE: 01-JUN-1995			
	CLASSIFICATION: 800			
	PRIOR APPLICATION DATA:			
	APPLICATION NUMBER: US 08/457, 205			
	FILING DATE: 01-JUN-1995			
	APPLICATION NUMBER: 08/258, 261			

US-09-929-769-7 (1-1029) x US-08-457-335A-6 (1-28958)

Oy	198	IlePheProLeuSerProAspProArgTrpGlnIleSerSerProAlaValAlaLeu	217
Db	26796	-----CCCCCTTGCTCTCCATCGAGGGCGCGTCTCTCTCCA---	26837
Oy	218	AlaLeuGlnGlnAlaLeuGlyGlnIleuLeuAlaArgValValGlnGly---SerProGlu	236
Db	26838	::: ::: ::: ::: ::: ::: :::	26887
Oy	237	ValProGlyITLeThrValAlaArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHis	256
Db	26888	ATGACGGCGGCACACGACGGCGCTTC-----AGCATGGCCCTGTGACTCGCCACGCCA----	26938
Oy	257	GlyGlyAlaLeuValMetSerMetHisArgSerHisPheLeuLacys-----	272
Db	26939	---CGCGTCTCTCCGGCAGAAGTTACCCTCC-TATTGCTTGATGATCTCCGGAATGC	26994
Oy	273	---ProLeuLeuArgGlnLeucysGlnTrpGlnArgcysValProGlnAspThr---Gly	290
Db	26995	CCCCC-----TCGGGCGGCACCCAGTGGACACCAAGT	27030
Oy	291	PheSerSerLeuPheLeuLysValLeuLeuGlnMetLeuGlnTrpLeuAspSerProGly	310
Db	27031	TTCCTACGCAGCC-----CTCGAGACCGCTCGG	27057
Oy	311	ValGlnGlyGlyPro-----LeuArgAlaGlnLeuArgMet	322
Db	27058	CTTGCTTAATGAGCCCGCAGTTCACAGGGCTCCGCCGCGCTCTACMAAGCGCGCAGCAGCT	27117
Oy	323	LeuAlaSerGlnAlaSerAlaGlyArgArg-----LeuSerAsp	335
Db	27118	CTTCGCCGGAAGCCAAAGCTCCCGGAGCCGCCGGAAGAGAGCGCGCTGTTCCTCCA	27177
Oy	336	ValArgGlyGlyLeuLeuArgLeuAlaGlnAlaLeuAlaPheArg-----	350
Db	27178	CCCCGCCCTGCTGCAGAGCGCTTCGACAGGCTCGCTTTGTAGACAGCACAGCAAAGC	27237
Oy	351	---GlnAspLeuGlnValAlaSerSerThrValArgAlaValIleAlaThrLeuArgSer	369
Db	27238	CTTAGAGATGCCCTTCTCGTAGAGCGAGT-----ATCGCTGCGCTCCGGTG	27285
Oy	370	GlyGlnGlnCysSerValGlnProAsp-----LeuIleSerLysValLeuGlnGly	386
Db	27286	-----GAGCACCAACCTCGCGGTTCGTTCCACCGTCGAGAGGG	27327
Oy	387	LeuITeGluValArgSerProHisLeuGlnIleuLeuThrAlaPheSerAlaThr	406
Db	27328	AATCTCGCGCTCGCTCC-----TCTCGCGC	27354
Oy	407	AlaSpAlaAlaSerProPheProLacysLys-----ProValValAlaSerSer	424
Db	27355	ACGCCAGAGCGCAACCATCGCTCGGTGCAAGCGCTCGCATCGCG-----	27402
Oy	425	LeuLeuLeuGlnGlnGlnIleuProLeuAlaGlyLysProGlyAlaAspGlyGlySer	444
Db	27403	-----CCGGTCCCGGACAGCAAGCTCCGACACCGGAGCG-----TCCAC	27444
Oy	445	LeuGlnAlaValaArgLeuGlyProSerSerGlyLeuLeuValaAspTrpLeuGlnMetLeu	464
Db	27445	CTGATGACCTC-----TTCGCAATCGACTGGACGAGCTGCA-	27482
Oy	465	AspProGluValaValser-----SerCysProAspLeuGlnLeuArgLeuLeu	480
Db	27483	AAGCCACCATCTACCCCATCGCCGCGAGGGGTCCCTCTCGGCACAGAGG-----	27536
Oy	481	PheSerArgArgGlyGlyGlnAlaGlnValProSerPheArgProTyLeuLeu	500
Db	27537	----TCTGACCTCGGACCAAGGT--GCTCTGACCCGATATACGACCTTGTCTCT	27590
Oy	501	Thr-----LeuPheThrHisGlnIleSerSerTrpProThrHisGln	514
Db	27591	ACGAGAGCCCTGCACACAGGCGTTGCGTCMAAGCCTGTATGCCCTTCATCGC	27650
Oy	515	CysITeArgValLeuLeuGlyLysSerArgGlnGln-----ArgPheAspPro	530

```

Db 27651 TCTGCCGAGGCGACCTCATTCGAGCGCCGAGACCCGCGCCGCTCGCCCT 27710
OY 531 SerAlaSerLeuSerPheLeuThrAlaCysAlaLeuHisValProArgIleTrpGlnGlyArg 550
Db 27711 CTTCGACAGCTG-----GCTCGCCGAGACGCGCCGCTCTCTCGCCGCTCGC 27758
OY 551 AspGlnArgThrProGlnIleArgArgGlnGlnLeuValLeuArgValGlnGlyProGlu 570
Db 27759 ---CTGCTGACCCGAGCGCCGCTCGCCGAGACCCGCTGAGAGAGAGCTCAAGGCGCTCG- 27814
OY 571 LeuIleSerLeuValGlnLeuIleLeuAlaGlnAlaGlnIleThrArgSerGlnAspGlyAsp 590
Db 27815 CTGACGCGCCCTCTGCGGCTGCTCGCTCGCTCGCCGAGACGAGACACCGAGCGCCCTC 27874
OY 591 ThrAlaIleSerLeuIleGlnAlaArgLeuProLeuLeuSerCysCysGly 610
Db 27875 TCGTCTCTGCGACCTCGACGACGAGCGCTCCGACGACGCGCTCTCGCGCGCTCG 27934
OY 611 AspArgGlnSerValArgGlyValThrGlnIleLeuSerGlyCysIleGlnGlnTrpGly 630
Db 27935 ACCGACAGAGAGCCAGAGATGCGCCCTCGCAAGGCAAGCCCTCGTTCGAA-----GGC 27988
OY 631 AspSerValLeuGlnIleArgArgCysArgAspLeuLeuLeuIleLeuTrpLeuGlnArgPro 650
Db 27989 TCTCAGCGCTGCCCCAGGCGCCGACG-----ACACAGCGTCC 28027
OY 651 GlnLeuArgValProValProGlnValLeuLeuHisSerGlnGlnAlaAlaSerSerSer 670
Db 28028 CCGCAGGCGCTCGGAGGAGCGCTG-CTCATCAGGAGGAGGACGCGCAGCGCTCGCGCGCTG 28086
OY 671 ValCysIleLeuAspGlyLeuIleHisArgPheIleThrLeuLeu-----685
Db 28087 TGTGCGCGCGCGCTCTGCTGAAACACAGCGCCAGCAGCCCTGCTCACCTCGCGCGCAG 28146
OY 686 -----AlaAspThrSerAspSerArgAlaLeuGlnAsnArgGlyAla 699
Db 28147 GCGCGCAGCGCTCGCGGTGATGTCCTTCCGAGGCGAG---CTCGAGAGCTCTGGGGGCT 28203
OY 700 AspAlaSerMetAlaCysArgGlyLeuAlaValAlaHisProLeuLeu---LeuLeuArg 718
Db 28204 TCGGTACACCTCGCGCGGTGCGAGCTGCGGAGTCCAGCGCTCTGAAGAGCATCTTGAT 28263
OY 719 HisLeuProMet-----IleAlaIleLeuLeuHisGlyArgThrHisLeuAsnPhe 735
Db 28264 AACATTTCGAGCGCTCACCGCGGTGCGCGCTGCTGATCCCGCAGCTCTCGAGCGC 28323
OY 736 GlnGluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGlnLeuGlu 755
Db 28324 GAT-----CTGCTCGCGCGCATAGC 28344
OY 756 LeuLeuGln---ProHisValPheArgSerGlnHisGlnGlnIleLeuTrpAspCysLeu 774
Db 28345 CTGACGCGGATGACCGCGCTCTCGCCGCCAAGATGATGATCGCGC---TGGCAC----- 28395
OY 775 LeuSerPheIleArgLeuLeuLeuAsnTrpArgGlySerSerArgHisLeuAlaIlePhe 794
Db 28396 -----TTCATCAGCTCACCCAGATAGACCCCTTGGCGCGCTTC 28434
OY 795 Ile-----AsnIlePheValGlnPheIleHisIleGlyTrpIleThrTrpAsnAlaProAla 812
Db 28435 ATGCTCTTCTCGTCGCTCGCGCGCGCTCTCGGACAGCTCAAGTCACTCACTCAAGCGCGCT 28494
OY 813 AlaIleSerPheLeuGlnIleHisAlaAsp-----ProLeuHis 825
Db 28495 GCGAGCGCGCTCTCTCGATGCGGTTCGCGCACACCGCGCGCGCAAGGCGCTCGCGCTCA 28554
OY 826 AspLeuSerPheAspAsn---SerAspLeuValMetLeuIleSerLeuLeuAlaGlyLeu 844
Db 28555 TCGCTGCGGTGAGCAGCTGCGCGCGAGCGCAGCGCAATGAGCAGACGACGCTCACCGCGCGC 28614
OY 845 SerLeuProSerArgAspArgTrpAspArg---GlyLeuAspGlnGlnGlnGlnGlu 863

```

```

Db 28615 GCGCGCCCT-----CGCATGAGAGCGCGCGCTTCCTCCCTCGACCTTGAGAG 28662
OY 864 GluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThr--- 882
Db 28663 AGG-----CTCGCCCTCTTGATGCGCGCGCTCTTCCGACCGAGACCGCC 28707
OY 883 -----AlaAlaGlnMetAlaPro 888
Db 28708 CTGATCCCGCGCGCTTCGACTTGAGCGCGCTCAGGCGCAAGCGCGCAGCTCCCGC 28767
OY 889 TyrMetIleArgLeuSerArgGlyGlnThrVal-----899
Db 28768 TTGTTCCACGCTCTCTGCGCGCTCGCAGCCGTACGCAAGGCGCGCAGCAACACCGCCAG 28827
OY 900 ---GluAspLeuLeuGlnValLeuSerAspIleAspGlnMetSerArgArgProGlu 918
Db 28828 GCCTCGCTGCTTACAGAGCGCGCTCTCAGCCCTCCCGCGCGCGCAAGCGCGCG---GCC 28884
OY 919 IleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSerAla 934
Db 28885 CTGCTCATGCTCATCCGACCGCAAGCGCGCGCGCTCTCGCGCTCGC 28932

RESULT 11
US-08-729-214-6
Sequence 6, Application US/08729214
Patent No. 5817502
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Hammett, Phillip E.
APPLICANT: Van Pee, Karl-Helinz
APPLICANT: Kirner, Sabine
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729, 214
FILING DATE: TBA
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,741
REFERENCE/DOCKET NUMBER: CGC 1506/CLIPS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-729-214-6

Alignment Scores:
Pred. No.: 0 006 Length: 28958
Score: 150.50 Matches: 255

```

Percent Similarity: 34.66% Conservative: 111
 Best Local Similarity: 24.15% Mismatches: 405
 Query Match: 2.87% Indels: 292
 DB: 1 Gaps: 57

US-09-929-769-7 (1-1029) x US-08-729-214-6 (1-28958)

QY 14 ThrLeuGIuProProArgAlaAspSerGluPheGlnAlaLeuAsp----- 30
 Db 26222 ACCTCCGGCAAAACCTCTCTGTTTTCGAGCGGAGACCGGCTCTCTGACATGGGCAAT 26280
 QY 31 IleTrpPheProGluGluLysProLeuPro-----ThrAlaPheLeuValAspThrSer 48
 Db 26281 CGCTCTCTCGAGAGTCAAGCCCATCGGCTGACAGCTCGGCTCGGCGAGACCTGC 26340
 QY 49 GluGluAlaLeuLeuLeuProAspTrpLeu---LysLeuArgMetLleArgSerGluVal 67
 Db 26341 GAGCGCTCAACCGCTGATCCGCTCGTGGCTCCATTTCAGCGAGCGAGACGACCACTC 26400
 QY 68 LeuArgLeuValAspAlaAlaLeuGlnAspLeuGluProGlnGlnLeuLeuPheVal 87
 Db 26401 GCCCGCTCTGCTC---CTCTCGTGGGCGAGCTCTTACCCGAGGCTCGGCTCGACCTGG 26457
 QY 88 GlnSerPheGlyIleProValSerSerMetSerLysLeuLeuGlnPheLeuAspGlnAla 107
 Db 26458 AACGCTCTTCGCGCGCTTCGCTCCCGCAGAGTCTCGCTCCGACCTAC----- 26508
 QY 108 ValAlaIleAspProGlnThrLeuGlnGlnAsnIleMetAspLysAsnTrpMetAlaHis 127
 Db 26509 -----CCCTCCCAACGCGAGCGCTTGTGGCTGACGCTCC---ACGCGGCGC 26553
 QY 128 LeuValGluValGlnHisGluArgGlyAlaSerGlyGlyGlnThrPhe-----HisSer 145
 Db 26554 GCTGCCGAGCTC-----GCTCCGAGCGCTTACCTCGGCGGACCAACCGC 26598
 QY 146 LeuLeuThrAlaSerLeuProProArgArgAsp----- 156
 Db 26599 CTGCTCGGCGCGCGCGCTGCG---CCTCGCGAGCGCGATGGCTTGTCTTCACAGAGCGCT 26657
 QY 157 ----- 158
 Db 26658 CTCCTCCGACGACGACCGCTGGCTCGAAGACAGCTGCTTCGGCATACCTGCTCTGC 26717
 QY 159 GluAlaProLysProLysSerSerProGlu---GlnProIleGlyGlnGlyArgIleArg 177
 Db 26718 CAGGCGCGCGCTCTCGAGCTGGCCCTGCATGCTCCGCCATCTCGTGGCTCGACACCT 26777
 QY 178 ValGlyThrGlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGln 197
 Db 26778 CGAAGCTCAACGCTCGA----- 26795
 QY 198 IlePheProLeuSerProAspProArgTrpGlnSerSerSerProArgProValAlaLeu 217
 Db 26796 -----CCCCCCCCCTGCTCTCCATCGCAGGCGCGCTCTCTCTCA--- 26837
 QY 218 AlaLeuGlnGlnAlaLeuGlyGlnGlnLeuAlaArgValAlaGlnGly---SerProGlu 236
 Db 26838 GATCTCCGTGGGCGCGGAGAGGTG---CTGGAC-----GAAGCGGCTCTCTCGTTC 26887
 QY 237 ValProGlyIleThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerProHis 256
 Db 26888 ATAGCGGCGCGACGACGCGCTTC-----AGGATGGCGCGCTGAGCTCGCACGCGCA--- 26938
 QY 257 GlyGlyAlaLeuValMetSerMetHisArgSerHisPheLeuAlaCys----- 272
 Db 26939 ---GGGCTCTCTCGCGAAGCTACCGCTCC---CATTCGCTTCGATGCTCCGCGAATGCG 26994
 QY 273 ---ProLeuLeuArgGlnLeuGlyGlnTrpGlnArgCysValProGlnAspThr---Gly 290
 Db 26995 CCCCC-----TCGGGCGCGCACCCAGCTGACACCAAGCT 27030
 QY 291 PheSerSerLeuPheLeuLysValLeuLeuGlnMetLeuGlnTrpLeuAspSerProGly 310
 ||| : : : ||| : : : |||

Db 27031 TTCACGCGAGCC-----CTCGAGACCGCTGGG 27057
 QY 311 ValGluGlyGlyPro-----LeuArgAlaGlnLeuLeuArgMet 322
 Db 27058 CTTGGCTTATGGCCGAGTTCAGGCGCTCGCGCGCTGTCTACAGCGGCGAGAGCT 27117
 QY 323 LeuAlaSerGlnAlaSerAlaGlyArg-----LeuSerAsp 335
 Db 27118 CTTCCGCCGAAGCCAGCTCCCGGAGCGCCGCCGAAGAGAGCGCGCTCTTGGCTCTCA 27177
 QY 336 ValArgGlyGlyLeuLeuArgLeuAlaGlnAlaLeuAlaPheArg----- 350
 Db 27178 CCCCCCGCTCTGACAGCGCTTGCAGGCGCTCGCTTGTAGAGACGAGCAAGAGCG 27237
 QY 351 ---GlnAspLeuGlnValValSerSerThrValArgAlaValIleAlaThrLeuArgSer 369
 Db 27238 CTTCAAGATGGCTTCTCTGAGCGGAGT-----ATCGCTCGCTCGGCTCG 27285
 QY 370 GlyGluGlnCysSerValGluProAsp-----LeuIleSerLysValLeuGlnGly 386
 Db 27286 -----GAGCACACACCTGCGGCTGGGCTTCCACGCTCGTCAAGAGCGG 27327
 QY 387 LeuIleGluValArgSerProHisLeuGlnGluLeuThrAlaPheSerAlaThr 406
 Db 27328 AATCTCGGCTCGCTCC-----TCCTCGCGG 27354
 QY 407 AlaAspAlaAlaSerProPheProAlaCysLys-----ProValValValValSerSer 424
 Db 27355 ACGCCAGAGCGAGCAACCATCGCTCGGTCAAGCGCTCGCGCATGGCGG----- 27402
 QY 425 LeuLeuLeuGlnGluGluProLeuAlaGlnGlyLysProGlyAlaAspGlyLys 444
 Db 27403 -----CCGGCTCGCGGAGACAGCTCCGAGACCGCGGAGCG-----TCCAC 27444
 QY 445 LeuGluAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrpLeuGluMetLeu 464
 Db 27445 CTTGAGTCCCTC-----TTCCGATGACAGCGAGCGAGCTGCA- 27482
 QY 465 AspProGluValValSer-----SerCysProAspLeuGlnLeuArgLeuLeu 480
 Db 27483 AAGCCCACTACGCGCATGCGCCCGGAGCGGCTGCTCTCGGACAGAAAG----- 27536
 QY 481 PheSerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTrpLeuLeu 500
 Db 27537 ---TTCGACACGCGGAGACAGGCT---GCTCTCGACCGCTATACGAGCTTGTCTCT 27590
 QY 501 Thr-----LeuPheThrHisGlnSerSerTrpProThrLeuHisGln 514
 Db 27591 ACGAGCGGCTCGACGAGGCGCTTGGCTCCCAACCTCGATACGCGCCCTTCATGCG 27650
 QY 515 CysIleArgValLeuLeuLysSerArgGluGln-----ArgPheAspPro 530
 Db 27651 TCTGCCCGAAGCGAGCTCATGCGAGCGCGCGGAGACCGCGGAGCGCTCGCCCT 27710
 QY 531 SerAlaSerLeuAspPheLeuTrpAlaCysIleHisValProArgIleTrpGlnArg 550
 Db 27711 CTTGCAAGCTG-----GCTCGCGACGAGCGCGCTCGCTCTCGGCGCTCGC 27758
 QY 551 AspGlnArgThrProGlnLysArgArgGluGluLeuValLeuArgValAlaGlnGlyProGlu 570
 Db 27759 ---CTCTCGCACCGCGCGCGCTGCGCACCGCTGAAGAGAGCTGAAGGCGCTCG- 27814
 QY 571 LeuIleSerLeuValGluLeuIleLeuAlaGlnAlaGlnIleThrArgSerLinspGlyAsp 590
 Db 27815 CTCAGCGGCTCTCTGGGCTCTCGCTCGCTCGCGCGAGAGGAGCAACCGAGCGGCTC 27874
 QY 591 ThrAlaAlaCysSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCysGly 610
 Db 27875 TCGTCTCTGTCACCTCGACGAGAGAGAGAGCTCCAGAGCGCGCTCGGCGGCTCG 27934
 QY 611 AspAspGluSerValArgLysValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGly 630
 Db 27935 ACGCAAGAGAGCGAGATCGCGCTCCGCAAGGCGCAACCGCTCGTTCCA-----GCG 27988
 ||| : : : ||| : : : |||

```

QY 631 AspSerValLeuGluArgArgCysArgAspLeuLeuGluLeuTyrLeuGluArgPro 650
DB 27989 TCTACAGCCTGCCCCAGAGCCGACCG-----ACACAGCGCTCC 28027
QY 651 GluLeuArgValProValProGluValLeuLeuHisSerGluGluAlaAlaSerSer 670
DB 28028 CCGCAGCCCTCGAGGACCGGTC-CTCATCAGCGGAGGACCGCAGCTCGGCGCCCTG 28086
QY 671 ValCysLysLeuAspLysLeuLeuHisArgPheLeuThrLeuLeu----- 685
DB 28087 GTGCGCGCGCGCTCGTGAACACAGAGCCAGACCTGCTCCACCTCGCGCCAG 28146
QY 686 -----AlaAspThrSerAspSerArgAlaLeuGluAsnArgGlyAla 699
DB 28147 GCGCGAGCGCTCCGGGTGTGTGCTTCGCGAAGCGAG---CTCGAAGCTCGGGGCT 28203
QY 700 AspAlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeu-----LeuLeuArg 718
DB 28204 TCGGTACACCTCGCCCGCGGCGGACGTCGCGGATCCAGCCGCTTAAGAGCTTCGTGAT 28263
QY 719 HisLeuProMet-----IleAlaAlaLeuLeuHisGlyArgThrHisLeuAsnPhe 735
DB 28264 AACATTCCAGCGCTCACCGCGTCCGCGCGCGCTGCGATCGCGCAGCGCTTCGAGCGC 28323
QY 736 GluGluPheArgGluGluHisHisLeuSerCysPheLeuHisValLeuGlyLeuGlu 755
DB 28324 GAT-----CTGCTCGCGCGCGCATGAGC 28344
QY 756 LeuLeuGlu-----ProHisValPheArgSerGluHisGluGlyAlaLeuTyrPaspCysLeu 774
DB 28345 CTGAGCGCGATCGACCGCGCTTCGCGCCCAACATGATGATCGCCG---TGGCAC----- 28395
QY 775 LeuSerPheLeuArgLeuLeuLeuAsnTyrArgLysSerArgHisLeuAlaAlaPhe 794
DB 28396 -----TTGCATCAGCTCACCACCAAGTAAGCCCTTGGCGGCTTC 28434
QY 795 Ile-----AsnLysPheValGluPheLeuHisLysTyrTrpLeuAsnAlaProAla 812
DB 28435 ATCTCTTCTCGCTCGCTCGCGCGCGCTCGGACGTCAGCTCAGCTCACTCACTCAGCGCGCT 28494
QY 813 AlaIleSerPheLeuGluHisAlaAsp-----ProLeuHis 825
DB 28495 GCGAGCGCTCTCGTGGCGCTGCGGACACCGCGCGCGGCGGAGGCGCTCGCTCGCTCA 28554
QY 826 AspLeuSerPheAspAsn-----SerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeu 844
DB 28555 TCGCTCGCGGAGCGACCTGCGCGGAGCGGACGCGCAATGACAGACGACGCGCGCGC 28614
QY 845 SerLeuProSerArgAspArgThrAspArg-----GlyLeuAspGluGluGlyGlu 863
DB 28615 GCGCGCGCT-----CGCATGGAGCGCGCGCGCTTCCCTCGACCTCTGAGGAG 28662
QY 864 GluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThr--- 882
DB 28663 AGG-----CTCGCGCTCTTTCGATGCGCGCTTTCGCAACCGAGACCGCC 28707
QY 883 -----AlaAlaGluMetAlaPro 888
DB 28708 CTGTCCTCGCGCGCTTCGACTGACCGCGCTAGGCGGAGCGGACCGCGCTCGCCCGC 28767
QY 889 TyrMetLysArgLeuSerArgGlyGluThrVal----- 899
DB 28768 TTGTTTCGAACGTTCTCGCGCTCGACCGCTGACGCAAGCGCGCGCAACACCGCCAG 28827
QY 900 ---GluAspLeuGluValLeuSerAspIleAspGluMetSerArgArgArgProGlu 918
DB 28828 GCGTCGCTCGCTTCAGAGCGCGCTTCAGCGCTCGCGCGCGGCGGAGCGGAGCT---GCC 28884
QY 919 IleLeuSerPhePheSerThrAsnLeuGluArgLeuMetSerSerAla 934
DB 28885 CTGCTCATCTCATCCGACACCGCGCGCGCTCTCGGCTCGCC 28932

```

```

RESULT 12
US-09-028-934-6
Sequence 6, Application US/09028934
Patent No. 6117670
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Hill, Dwight S.
APPLICANT: Lam, Steven T.
APPLICANT: Hammer, Philip E.
APPLICANT: van Pee, Karl-Heinz
APPLICANT: Kirner, Sabine
APPLICANT: Young, Thomas R.
TITLE OF INVENTION: Pyroglutinin Biosynthesis Genes and Uses
TITLE OF INVENTION: Theoret
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6117670artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/028,934
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC1506/CIP7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: Linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-028-934-6
Alignment Scores:
Pred. No.: 0.006 Length: 28958
Score: 150.50 Matches: 255
Percent Similarity: 34.66% Conserved: 111
Best Local Similarity: 24.15% Mismatches: 405
Query Match: 2.87% Indels: 292
DB: 3 Gaps: 57
US-09-929-769-7 (1-1029) x US-09-028-934-6 (1-28958)
QY 14 ThrLeuGlyProProArgAlaAspAspSerGluPheGluAlaLeuLeuAsp----- 30
DB 26222 ACCTCGCGAACCCTGCTTC-TCGAGCGGACGAGCGGCTCTCTCGACGATGGGCAT 26280
QY 31 IleTrpPheProGluGluLysProLeuPro-----ThrAlaPheLeuValAspThrSer 48
DB 26281 CGCTTCTTCGAGGTCACGCGCCCATCCGTCGTCACGCTCGCGCGCGGAGACTCG 26340

```


OY	49	gluglAlaLeuLeuLeuProAspIrrleu---	LysLeuArgMetLleArgSerLVal	67
Db	26341	GAGCGCTACGCGTGGATCCGTCTGTGGCTCCATTGCAGCGCAGCAAGCACACCCTC	...	26400
OY	68	LeuAglLeuValAspAlaLeuInAspLeuGluproGlnGluLeuLeuPheVal	87	
Db	26401	GCGCGCTGTC---CTCTCCCTGGGGCGGAGCTCTTACCAGAGGCTCGCGCTGCACATCG	26457	
OY	88	GlnSerPheGlyIleProValSerSerMetSerLysLeuLeuGlnPheLeuAspGlnAla	107	
Db	26458	AACGCCCTTTTGCGCCCTTCCTCCGCGCAGAGCTCCCTCCACCACTAC-----	26508	
OY	108	ValAlaHisAspProGlnThrLeuGlnGlnAsnIleMetAspLysAsnTyrMetLahis	127	
Db	26509	-----CCCTCCACAGCAGAGGGCTTCTGGCTTCGAGCGCTCC---ACGGCGAC	26553	
OY	128	LeuValGluValGlnHisgLuArgylalaserGlyGlnThrPhe-----Hisser	145	
Db	26554	GCTGCCGACGTC-----GCCTCCGAGGCTTACCTCGCGCCACCAACCG	26598	
OY	146	LeuLeuThrAlaSerLeuProProArgArgasp-----	156	
Db	26599	CTGCTCGGGCGCGCGCTGC-CCTCGCGACCGCATGGCTTGTCTTCACAGACGGCT	26657	
OY	157	-----Serrthr	158	
Db	26658	CTCCCTCGCAGACACCGCTGCTCGAAGACCAGTCGTTCGGCATCCCTGTCTGC	26717	
OY	159	GluAlaProLysProLysSerSerProGlu---GlnProIleGlyGlnArgIleArg	177	
Db	26718	CAGGGCGCGCTCTCTGACTGCGCTGCATGTGCGCCCTCTCGTGGGCTTGACACCGT	26777	
OY	178	ValGlyThrGlnLeuArgValLeuGlyProGluAspLeuAlaGlyMetPheLeuGln	197	
Db	26778	CGAAGACGTCACGCTCGA-----	26795	
OY	198	IlePheProLeuSerProAspProArgTyrGlnSerSerSerProArgProValAlaLeu	217	
Db	26796	-----CCCCCCCCTGCTCCCATCGACAGGGCGCGCTCTCTCTCCA--	26837	
OY	218	AlaLeuGlnGlnAlaLeuGlyGlnGluLeuAlaIleArgValAlaGlnGly---SerProGlu	236	
Db	26838	GATCTTCGTCGGGGCGCGGACGGTG-CTGGAC-----GAAGGGCGCTCTCCGTTTC	26887	
OY	237	ValProGlyIleThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProhis	256	
Db	26888	ATAGCGGGCGGCACAGCGGGCTTC-----ACAGTAGGCCCCCTGGACCTGCACAGCCA--	26938	
OY	257	GlyGlyAlaLeuValMetSerMetHisArgSerHisPheLeuAlaCys-----	272	
Db	26939	---GGCGCTCTTCGCGCAGACTACCGCTCC-CATTGCTCTCGATGCCCGCATGTC	26994	
OY	273	---ProLeuLeuArgGlnLeuCysGlnTyrGlnArgCysValProGlnAspThr---Gly	290	
Db	26995	CCCCC-----TCGGGCGCCACCCAGGTGACACCCAAAGT	27030	
OY	291	PheSerSerLeuPheLeuLysValLeuLeuGlnMetLeuGlnTrpLeuAspSerProgly	310	
Db	27031	TTCACGCGACG---CTCGAGAGCGCTGG	27057	
OY	311	ValGluGlyGlyPro-----LeuArgAlaGlnLeuArgmet	322	
Db	27058	CTTGCTTAATGGCCGAGTCCAGGCGCTCCGCGCGCTCTCAAGCGGGCGAGAGCT	27117	
OY	323	LeuAlaSerGlnAlaSerAlaGlyArgarg-----LeuSerasp	335	
Db	27118	CTTCGCCGGAAGCAAAGCTCCGAGCGCGCCGACAGAGACGCCGCTGTTTTGCCCTCCA	27177	
OY	336	ValArgGlyGlyLeuLeuArgLeuAlaGlnAlaLeuAlaPhearg-----	350	
Db	27178	CCCGCGCGCTCGACAGAGCGCTTCGAGCGCTGTCGCTTTGTAGAACAGACAGCAAGG	27237	
OY	351	---GlnAspLeuGluValValSerSerThrValArgAlaValIleAlaThrLeuArgSer	369	

Db	27238	CTTCAGAGATGCCCTTCTCTGTGACGGAGT	-----ATGCGTCTGCTCCGGCTG	27285
QY	370	gLygluInCysSerValgluProasp	-----LeuIleSerLysValLeuIngly	386
Db	27286	-----GAGCCACACCCCTCGGGTGGTTTCACACGCTCTGAGGGCG	-----	27327
QY	387	LeuIleGluValArgSerProHisLeuIngluLeuLeuThrAlaPhePheSerAlaThr	-----	406
Db	27328	AATCTCGCGCTCGCTCC	-----TCTCTCGCG	27354
QY	407	AlaAspAlaIleSerProPheProAlaCysLys	-----ProValValValSerSer	424
Db	27355	ACGCCAGAGCGCAACCCATCGCTCGGTGCACAGCCCTGCCATGCGCG	-----	27402
QY	425	LeuLeuLeuIngluIngluProLeuAlaGlyGlyLysProGluValAlaAspIleGlySer	-----	444
Db	27403	-----CCGCGTCCGCCGAGACAGCTCCGACACCCGGAGCG	-----TCCAC	27444
QY	445	LeuGluAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTyrLeuIleuLeu	-----	464
Db	27445	CTCGATGCCCTC	-----TTCCGATCGACGTGAGCGACGTGCA	27482
QY	465	AspProGluValValSer	-----SerCysProAspLeuInleuArgLeuLeu	480
Db	27483	AAGCCCACTACCGCCATCGCCCGACGGGTGCCCTCTCGGCACAGAGG	-----	27536
QY	481	PheSerArgArgGlyGlyGlyGluInAlaProSerPheArgProTyrLeuLeu	-----	500
Db	27537	-----TCTCGACCTCGGACCAAGGT	-----GCTCTCGACCCCTTACCGACTTGTCTCT	27590
QY	501	Thr	-----LeuPheThrHisGlnSerSerTyrProThrLeuHisGln	514
Db	27591	ACGACAGCGCTTCGACACAGCGGCTTCGCCCAACCTCGTCATCGCCCTTCATCGC	-----	27650
QY	515	CysIleArgValLeuLeuGlyLysSerArgGluIn	-----ArgPheAspPro	530
Db	27651	TCGCCCCGAGGCGACCTCATCGACGAGCCCGCGAGACCAACCGCGACGCTCGCTG	-----	27710
QY	531	SerAlaSerLeuAspPheLeuTyrAlaCysIleHisValProArgIleTyrGlnLeuArg	-----	550
Db	27711	CTTCGACGCTG	-----GCTCGCGACAGCGGCTCGCTCGCTCGCTCGCTCG	27758
QY	551	AspGlnArgTyrProGlnLysArgArgIngluLeuValLeuArgValIngluProGlu	-----	570
Db	27759	-----CCTCGTCACCCGACGCGCGCTCGCGCCACCCAGCTGAAGAAAGTCACAGGCGCTCG	-----	27814
QY	571	LeuIleSerLeuValGluLeuLeuIleLeuAlaGluAlaInThrArgSerGlnAspIleAsp	-----	590
Db	27815	CTCACGCGCTTCGCGGTCTCGCTCGCTCGCGCGACGACGACCAACCGAGCGGCTCG	-----	27874
QY	591	ThrAlaAlaCysSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCysGly	-----	610
Db	27875	TCGTCTCTCGACCTCGACGACGACGAGGAGGCTCCAGCAGCCCTGCTGGCGCGCTCG	-----	27934
QY	611	AspAspGluSerValArgLysValThrGlnHisLeuSerGlyCysIleGlnInTyrGly	-----	630
Db	27935	ACGCAAGAGACGACGAGTCGCGCTCGCGCAACGCGCAACCCCTCGTCCAA	-----GCG	27988
QY	631	AspSerValLeuGlyArgArgCysArgAspLeuLeuGlnLeuTyrLeuInArgPro	-----	650
Db	27989	TCCTACGCTGCCCGACGGCGGCCACGG	-----ACACAGCGTCC	28027
QY	651	GluLeuArgValProValProGluValLeuLeuHisSerGluGlyAlaAlaSerSerSer	-----	670
Db	28028	CCGCAAGGCTCGGAGGACCGCTC	-----CTCATCAGCGGAGGACCGGACGCTCGGCGCCCTG	28066
QY	671	ValCysLysLeuAspIleLeuIleHisArgPheIleThrLeuLeu	-----	685
Db	28087	GTCGCGCGCGCTCGTCTAAACACGACGCGCAAGACACTGCTCTCACTCGCGCCAG	-----	28146
QY	686	-----AlaAspThrAspSerArgAlaLeuGlnAsnArgGlyAla	-----	699

```

Db 28147 GGGCGAGCGCTCGGCTGTATGCTTCGAGAGAG---CTCGAGCTTGGGGGCT 28203
QY 700 AspAlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeu---LeuLeuArg 718
Db 28204 TCGGTACACCTCGCGGTGCGAGCGTGGCGAGTCCAGCGCGCTTAAGAGACCTTCGAT 28263
QY 719 HisLeuProMet-----IleAlaAlaLeuLeuHisGlyArgThrHisLeuAsn 735
Db 28264 AACATTCCGAGCGCTACCGCGGTGGCGCGCTGTGATCGCGCGCTCGAGAGG 28323
QY 736 GluGluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGluLeuGlu 755
Db 28324 GAT-----CTGCTCGCGCGCATAGC 28344
QY 756 LeuLeuGln---ProHisValPheArgSerGlnHisGlnIleValLeuTrpAspCysLeu 774
Db 28345 CTCGACGCGATGACCGCGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 28395
QY 775 LeuSerPheIleArgLeuLeuLeuAsnTyrArgLysSerSerArgHisIleAlaPhe 794
Db 28396 -----TTGCATCAGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 28434
QY 795 Ile-----AsnLysPheValGlnPheIleHisLysTyrIleThrTyrAsnAlaProAla 812
Db 28435 ATCCTCTTCGTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 28494
QY 813 AlaIleSerPheLeuGlnLysHisAlaAsp-----ProLeuHis 825
Db 28495 GCGAGCGCGCTTCCTCGATGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCA 28554
QY 826 AspLeuSerPheAsn---SerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeu 844
Db 28555 TCGCTGCGGTGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 28614
QY 845 SerLeuProSerArgAspAspArgThrAspArg---GlyLeuAspGlnGlnGluGlu 863
Db 28615 GCGCGCGCT-----CGCATGAGCGCGCGCGCGCGCGCTCGCGCGCGCGCGAG 28662
QY 864 GluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThr 882
Db 28663 AGG-----CTCGCGCTTCGATGCGCGCGCTTCGCGCGCGCGCGCGCGCGCG 28707
QY 883 -----AlaAlaGluMetAlaPro 888
Db 28708 CTGCTCGCGCGCGCGCTTCGATGAGCGCGCTGAGCGCGCGCGCGCGCGCGCGCG 28767
QY 889 TyrMetLysArgLeuSerArgGlyGlnThrVal-----899
Db 28768 TTGTTCCAGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 28827
QY 900 ---GluAspLeuGlnValLeuSerAspIleAspGlnMetSerArgArgProGlu 918
Db 28828 GCGCTCGCTTCAGAGCGCGCTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 28884
QY 919 IleLeuSerPheSerThrAsnLeuGlnArgLeuMetSerAla 934
Db 28885 CTGCTGATGTATCCGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 28932

```

```

; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,118
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
;
; US-08-993-118-1
;
; Alignment Scores:
; Pred. No.: 0.000359 Length: 4031
; Score: 147.00 Matches: 203
; Percent Similarity: 33.80% Conservative: 113
; Best Local Similarity: 21.71% Mismatches: 312
; Query Match: 2.80% Indels: 310
; Gaps: 38
;
; US-09-929-769-7 (1-1029) x US-08-993-118-1 (1-4031)
;
; QY 86 PheValGlnSer---PheGlyIleProValSerSerMetSerLys-----LeuLeuGln 102
; Db 728 TTTCGCCAGTCTCTCTCCAGATTCCTGTGAGCTCTCTCTCTCTCTCTCTCTCTCT 787
; QY 103 PheLeuAspGlnAlaValAlaHisAsp-----ProGlnThrLeu 115
; Db 788 CTTCACAGATTCCTCTGAGAGTCAAGATCTTTGAGGATTTTCCCGAGTCTCTCC 847
; QY 116 GluGlnAsnIleMetAspLysAsnTyrMetAlaHisLeuValGluValGlnHisGluArg 135
; Db 848 CTCACAGATTCCTATGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 907
; QY 136 GlyAlaSerGlyGlnThrPheHisSerLeu-----146
; Db 908 CCTGAGAGTGTCAAGATCTTTGAGGATTTTCCCGAGTCTCTCTCCAGATTCCTGCG 967
; QY 147 -----LeuThrAlaSerLeuProProArgArgAsp 156
; Db 968 AGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1027
; QY 157 SerThrGluAlaProLysProLysSerSerProGlnGlnProIleGlyGlnArgGlyIle 176
; Db 1028 AGTACTTTTGAGGTTTCCCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1081
; QY 177 ArgValGlyThrGlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeu 196
; Db 1082 TTCTCTCTACTTATGAGATTTTCCAGAGTTCTCTCTGAGAGTGTCAAGTACTTTT 1141
; QY 197 GlnIlePheProLeuSerPro-----AspProArgTyrPglIns 209
; Db 1142 GAGGTTTTCGCCAGTCTCTCTCCAGATTCCTGAGAGCGCGCGCGCGCGCGCGCGCG 1201

```

[illegible]

Db	2013	TCGAGTCTCTCTCCATAGTC-----	2033
Qy	555	ProGlnLysArgArgGlnGluLeuValLeuArgValGlnGlyProGluLeuIleSerLeu	574
Db	2034	-----CTCAGAGCCCTCTCGAGGGGATGCATCTCCAAATCTCTCTCCAGA	2078
Qy	575	ValGluLeuIleLeuAlaGluAlaGluThrArgSerGlnAspGlyAspThrAlaIacys	594
Db	2079	GTCCTGAGAGG-----	2090
Qy	595	SerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCysGlyAspAspIuSer	614
Db	2091	-----CTCCTGAGGGGAGAGATCTCCCTGTCTCTCTCC-----AAATTC	2129
Qy	615	ValArgLysValThrGlnIleLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeu	634
Db	2130	CTCAGAGCTCTCTTGAGGGAGAGACTCCCTGT-----	2162
Qy	635	GlyArgArgCysArgAspLeuLeuGlnLeuThrLeuGlnArgProGluLeuArgVal	654
Db	2163	-----CTTCTCTCATTTTCTCAGA-----	2183
Qy	655	ProValProGluValLeuLeuIleHisSerGlnGlyAlaAlaSerSerSerValCysLysLeu	674
Db	2184	-----GTCTCTCTGAGTGGG-----AGCATC	2204
Qy	675	AspGlyLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeu	694
Db	2205	CCCTCTCTCTCTCCACTTTCTCCAGTTCTCTCTC-----	2240
Qy	695	GluAsnArgGlyAlaAspAlaSerMetAlaCysArgGlyLeuAlaValAlaHisProLeu	714
Db	2241	-----AGGGGAGAGACTCCAGACTCTCTCTCCAGAGCTCGTAGATCTCTCTCT	2294
Qy	715	LeuLeuLeuArgHisLeuProMetIleAla-----AlaLeuLeuHisLysL	729
Db	2295	CCACTCTCTTTAGTCTTCCCAAGATTCTCTCCAGAGCTCCAGAGCTCTCTCTGGGGG	2354
Qy	730	ArgThrIleLeuAsnPheGlnGlnThrPheArgGlnGlnAlaHisLeuSerCysPheLeuHis	749
Db	2355	CTGCTCACTCTCCCTCCAGAGAC-----CTGTCAAGCTCC-TTCTCTC	2395
Qy	750	ValLeuGlyLeuLeuGlnLeuGlnProHisValPheArgSerGluHis-----	766
Db	2396	TCCTACACTTTAGGAGACTCTTCTCA-----AGTTCCATAGAGCTCT	2440
Qy	767	-----GlnGlyAlaLeuThrAspCysLeuLeuSerPheIle	778
Db	2441	CAGAGCTCTCTGAGGGGCTGCCCAAGTCTCTCTCCAGAGCTCTGTAGCTCTCTTCCC	2500
Qy	779	ArgLeuLeuLeuAsnTyrArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPhe	798
Db	2501	TCCTTCCACTTCATGAGACTTTTCCCAAGATTCTCTGTAGCTCTCTCCCTCCACT	2560
Qy	799	ValGlnPheIleHisLys-----TyrIlePheThr	808
Db	2561	TCATGAGCTCTTCCAGAGCTCCCTCCAGAGCTCTCTCCAGAGCTCTGTGTATCTCTTC	2620
Qy	809	AsnAlaProAlaAlaIleSer-----PheLeuGlnLysHisAlaAspProLeuHisAspLeu	827
Db	2621	TCCTCTCCACTTCATTTAGAGCCCATTTCACTGAGAGATCCAGAGGCCCAAGTACATGATAT	2680
Qy	828	SerPheAspAsnSerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeuPro	847
Db	2681	ACAAGTTCCTCCAGACACCTCTAGAGAGTATCTTCTACAGACAGCAGACGATCTTGATA	2740
Qy	848	SerArgAspAspArgThrAspArgGlyLeuAspGlnGlnGlyGlnGlnLysIleSerSerAla	867
Db	2741	GAGAGCCAGCCCTTTGCTCACTTTATACACTGTGATGAAGAAAGCTGGACAG-----TTG	2791
Qy	868	GlySerLeuProLeuValSerValSerLeuPheThrProLeuThrAlaAlaGluMetAla	887

Db 2792 GCCGGTTCCTCTCTCAATATCAAGTACAGCCCTATCAAAAGCAGATGCTG 2851
 Qy 888 -----ProlymethylsargLeuSerArgGly 896
 Db 2852 ACAGATGTCATCAGACGATGACGGGCTACTTCTGTGATCTTCAGAAAGCCCGT--- 2908
 Qy 897 GlnThrValGlnAspLeuLeuGluValLeuSerAspIle 909
 Db 2909 -----GAGTTCATAGAGATCTTTTGGCAT 2935
 RESULT 14
 US-08-845-528C-1
 ; Sequence 1, Application US/08845528C
 ; Patent No. 6027924
 ; GENERAL INFORMATION:
 ; APPLICANT: LUCAS, Sophie;
 ; APPLICANT: DE SMET, Charles;
 ; APPLICANT: BOON-FALLEUR, Thierry;
 ; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
 ; TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pelfe & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York City
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10022
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/845,528C
 ; FILING DATE: April 25, 1997
 ; CLASSIFICATION: 4335
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mary Anne Schofield
 ; REGISTRATION NUMBER: 36,669
 ; REFERENCE/DOCKET NUMBER: LUD 5455
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 688-9200
 ; TELEFAX: (212) 838-3884
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4031 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double-stranded
 ; TOPOLOGY: linear
 ; US-08-845-528C-1
 Alignment Scores:
 Pred. No.: 0.000359 Length: 4031
 Score: 147.00 Matches: 203
 Percent Similarity: 33.808 Conservative: 113
 Best Local Similarity: 21.71% Mismatches: 312
 Query Match: 2.80% Indels: 310
 DB: 3 Gaps: 38
 US-09-929-769-7 (1-1029) x US-08-845-528C-1 (1-4031)
 Qy 86 PheValGlnSer---PheGlyIleProValSerSerMetSerLys-----LeuLeuGln 102
 Db 728 TTTCCCAAGTCTCTCTCCAGATTCTGTGAGCTCTCTCCCTCCCACTTTATTGAGT 787
 Qy 103 PheLeuAspGlnAlaValAlaHisAsp-----ProGlnThrLeu 115
 Db 788 CTTTTCACAGATGCTCCCTGAGACACTGAGTCTTTGAGGGTTCCTCCCAAGTCTCTT 847
 Qy 116 GlnGlnAsnIleMetAspLysAsnIlyMetAlaHisLeuValGlnIleHisGluArg 135

Db 848 CTCACATTCCTTTGACCT 907
 Qy 136 GlnAlaSerGlyGlnThrPheHisSerLeu----- 146
 Db 908 CCTGAGAGTCTCTCAAGTACTTTTGAAGGTTTCCCAAGTCTCTCTCCAGATTCTGGG 967
 Qy 147 -----LeuThrAlaSerLeuProProAlaGlyAsp 156
 Db 968 AGCCCT 1027
 Qy 157 SerThrGlnAlaProLysProLysSerSerProGlnIleProIleGlnIleArgGly 176
 Db 1028 AGTACTTTTGAAGGTTTTCCTCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1081
 Qy 177 ArgValGlyThrGlnLeuArgValLeuGlyProGlnAspAspLeuAlaGlyMetPheLeu 196
 Db 1082 TTCT 1141
 Qy 197 GlnIlePheProLeuSerPro-----AspProArgThrGlnSer----- 209
 Db 1142 GAGGGTTTCCCAAGTCT 1201
 Qy 210 -----SerSerProArg----- 213
 Db 1202 CTGAGCTTTTCCAGAGTCT 1261
 Qy 214 ---ProValAlaLeuAlaLeuGlnAlaLeuGlnIleValAlaArgValAlaGln 232
 Db 1262 TCT 1321
 Qy 233 GlySerProGln-ValProGlyIleThrValAlaArgValLeuGlnAlaLeuAlaThrLeu 252
 Db 1322 AGTTCTCTCTGAGAGTCT 1373
 Qy 252 userSerProHisGlyGlyAlaLeuValMetSerMetHisArgSerHisPheLeuLac 272
 Db 1374 -----TTCAGATCTCTGTGAGC 1390
 Qy 272 sProLeu-----LeuArgGlnLeuGlySer 280
 Db 1391 TCT 1450
 Qy 280 nIlyGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeu-LysValLeu 300
 Db 1451 ACTTTGAGGGTTTCT 1501
 Qy 300 euGlnMetLeuGlnIlePheLeuAspSer---ProGlyValGlnGlyGlyProLeuArgAlaG 319
 Db 1502 TCT 1560
 Qy 319 ILeuArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyG 339
 Db 1561 TTTTGAAGGTTTCCCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1601
 Qy 339 ILeuLeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGlnValValSerSer 359
 Db 1602 -----GTCCTCTCTGAGGGAGGAAATCCCATTTCTCTCTCTCTCTCTCTCTCT 1638
 Qy 359 hrValArgAlaValIleAlaThrLeuArgSerGlyGlnGlnCysSerValGlnProAsp 379
 Db 1639 GATTGTTCCAGTCT 1677
 Qy 379 euIleSerLysValLeuGlnGlyLeuIleGlnValArgSerProHisLeuGlnIleLeu 399
 Db 1678 TCT 1728
 Qy 399 euThrAlaPhePheSerAla-ThrAlaAspAlaAlaSerProPheProAlaCysLysPro 418
 Db 1729 TCACTACTTCT 1770
 Qy 419 ValValValValSerSerLeuLeuGlnGlnGlnIleGlnIleProLeuAlaGlyGlyLysPro 438
 Db 1771 -----GTCCT 1815

QY 578 eleuAlaGluAlaGluThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleG1 598
Db 2300 G-----CCTCTAG 2308
QY 598 nAlaArgLeuProLeuLeuLeuSerCysCysGlyAspAspGluSerValArgGlySva 618
Db 2309 GGGAGGATTCCTCTGCTCTCC-----AAATTCCTCAGAGTCT 2350
QY 618 lThGlnHisLeuSerGlyCysIleGlnGlnTProGlyAspSerValLeuGlyArgGlyCy 638
Db 2351 CTTCAGGAGAGAGACTCCCTGT----- 2372
QY 638 sArgAspLeuLeuGlnLeuTyrLeuGlnArgProGluLeuArgValProValProG1 658
Db 2373 -----CTCTCTCCATTTCCTCAGA----- 2393
QY 658 uValLeuLeuHisSerGlnGlyAlaAlaSerSerSerValCysIlySLeuAspGlyLeuI1 678
Db 2394 -GTCTCTCTGAGTGG-----AGGACTCCCTCTCTCT 2425
QY 678 eHisArgPheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGluAsnArgI 698
Db 2426 CTCCACTTTCCTCAGATTTCCTCTC-----AGGGG 2455
QY 698 yAlaAspAlaSerMetAlaCysArgGlyLeuAlaValAlaHisProLeuLeuLeuAr 718
Db 2456 GAGGACTTCAGTCTCTCTCTCCAGAGTCTGTGAGTATCTGTCTCTCTCCACTTCTTG 2515
QY 718 gHisLeuProMetIleAla-----AlaLeuLeuHisGlyArgThrHisLe 733
Db 2516 AGTCTTCCCTCAGATTTCCTGAGAGTCTCAGAGTCTCCTGAGGGGCTCTCAGTCT 2575
QY 733 uAsnPheGlnIleuPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGlyLe 753
Db 2576 CCTCTCCAGAGAC-----CTGTACGCTCC-TCCTTCTCTACACTTT 2616
QY 753 uLeuGluLeuGlnProHisValPheArgSerGlnHis----- 766
Db 2617 AGCGAGTCTCTCCAA-----AGTCCCATGAGAGTCTCCAGAGTCTCTCC 2661
QY 767 -----GlnGlyAlaLeuTyrAspCysLeuLeuSerPheIleArgLeuLeuLe 782
Db 2662 TGAGGGGCTGCTCCCTCTCCAGAGTCTCTGAGTCTCTCCCTCTCCACTTC 2721
QY 782 uAsnTyrArgIysSerSerArgHisLeuAlaAlaPheIleAsnIysPheValGlnPheI1 802
Db 2722 ATCGAGTCTTCCAGAGTCTCTCTGAGTCTCTCCCTCTCCACTTCATCGAGTCT 2781
QY 802 eHisIys-----TyrIleThrTyrAsnAlaProAl 812
Db 2782 TTCCAAGAGTTCCTGAGAGTCTCTCCAGAGTCTGTGATCTCTCTCTCTCCAC 2841
QY 812 aAlaIleSer---PheLeuGlnIysHisAlaAspProLeuHisAspLeuSerPheAspAs 831
Db 2842 TTCATTGAGCCCATTCAGTGAAGAGTCCAGCCAGCCAGTAGATGAATATACAAAGTTCTC 2901
QY 831 nSerAspLeuValMetLeuIysSerLeuLeuAlaGlyLeuSerLeuProSerArgAspAs 851
Db 2902 AGACACCTTGTGAGAGTGTATCTTGACAGACAGCCAGTCTTGATAGAGAGCGAGCC 2961
QY 851 PArgThrAspArgIlyLeuAspGlnGlnGlyLeuGlnIysSerSerAlaGlySerLeuPr 871
Db 2962 CTGTGTCACCTATACACTGATGAAGAGTGGAGAG-----TTGGCGCGGTTTCT 3012
QY 871 oLeuValSerValSerLeuPheThrProLeuThrAlaAlaGluMetAla----- 887
Db 3013 TCTCTCAAAATATCAAGTGAAGCAGCTATCACAAGAGCAGAGATGACGAATGTCAT 3072
QY 888 -----ProTyrMetIysArgLeuSerArgGlyGlnThrValG1 900
Db 3073 CAGCAGGTACACGGGCTTCTTCTGTGATCTTCAGGAAGCCCGT----- 3118

QY 900 uAspLeuGlnuValLeuSerAspIle 909
Db 3119 -GAGTTCATAGAGATACTTTTGGCAT 3145

Search completed: May 14, 2003, 02:39:10
Job time : 574 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 13, 2003, 23:33:17 ; Search time 314 Seconds
(without alignments)
4072.431 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 5241
Sequence: 1 MHTLVVHAWVLTLLGPPRA.....MDSQAISEALRIHMEAVM 1029

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 783854 seqs, 621352466 residues
Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEY=xjh
-MODEL=frame+p2n.model -DB=us09-929-769-7/us09-929-769-7.p2n.rnpb -MINMATCH=0.1
-O=us09-929-769-7/us09-929-769-7.p2n.rnpb -START=1 -END=1 -MATRIX=blomsu62
-DB=us09-929-769-7/us09-929-769-7.p2n.rnpb -START=1 -END=1 -MATRIX=blomsu62
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62
-TRANS=us09-929-769-7/us09-929-769-7.p2n.rnpb -SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=us09929769.ecgn1.1.189 -runat.07052003.161926.8130
-NCPU=6 -ICPU=3 -NO_XMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

- 1: /cgnt2_6/ptoddata/2/pubpna/us07_PUBCOMB.seq:*
- 2: /cgnt2_6/ptoddata/2/pubpna/PCIT_NEW_PUB.seq:*
- 3: /cgnt2_6/ptoddata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgnt2_6/ptoddata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgnt2_6/ptoddata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgnt2_6/ptoddata/2/pubpna/PCITUS_PUBCOMB.seq:*
- 7: /cgnt2_6/ptoddata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgnt2_6/ptoddata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgnt2_6/ptoddata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgnt2_6/ptoddata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgnt2_6/ptoddata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgnt2_6/ptoddata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgnt2_6/ptoddata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgnt2_6/ptoddata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	DB ID	Description
1	5241	100.0	3437	9 US-10-033-245-21
2	5241	100.0	3437	9 US-10-033-223-21
3	5241	100.0	3437	9 US-10-033-167-21
4	5241	100.0	3437	9 US-10-033-244-21

5	5241	100.0	3437	9 US-10-033-435-21	Sequence 21, Appl
6	5241	100.0	3437	9 US-10-032-990-21	Sequence 21, Appl
7	5241	100.0	3437	9 US-09-929-769-3	Sequence 3, Appl
8	5241	100.0	3437	9 US-10-033-996-21	Sequence 21, Appl
9	5241	100.0	3437	9 US-10-033-386-21	Sequence 21, Appl
10	5241	100.0	3437	12 US-10-033-246-21	Sequence 21, Appl
11	5241	100.0	3437	12 US-10-033-301-21	Sequence 21, Appl
12	5241	100.0	3437	12 US-10-033-326-21	Sequence 21, Appl
13	4211	80.3	2824	10 US-09-822-849A-276	Sequence 276, App
14	2055	39.2	1559	9 US-09-984-271-56	Sequence 56, Appl
15	924	17.6	639	9 US-10-040-739-1131	Sequence 1131, Ap
16	708	13.5	406	9 US-09-918-995-2971	Sequence 2971, Ap
17	592	11.3	484	9 US-09-764-891-1468	Sequence 1468, Ap
18	570	10.9	443	10 US-09-963-965-4270	Sequence 4270, Ap
19	565	10.8	423	10 US-09-983-965-3822	Sequence 3822, Ap
20	374.5	7.1	6413	9 US-10-108-605-26	Sequence 26, Appl
21	317	6.0	436	10 US-09-783-590-9285	Sequence 9285, Ap
22	163.5	3.1	14800	10 US-09-954-456-1601	Sequence 1601, Ap
23	157.5	3.0	7419	10 US-09-815-242-4009	Sequence 4009, Ap
24	152.5	2.9	4080	10 US-09-962-436-342	Sequence 342, App
25	151.5	2.9	3654	9 US-09-938-842A-1225	Sequence 1225, Ap
26	151.5	2.9	7015	9 US-10-132-382-5	Sequence 5, Appl
27	151.5	2.9	7093	9 US-10-132-382-1	Sequence 1, Appl
28	151.5	2.9	7229	9 US-10-132-382-7	Sequence 7, Appl
29	151.5	2.9	7307	9 US-10-132-382-3	Sequence 3, Appl
30	147	2.8	4031	9 US-10-085-108-1	Sequence 1, Appl
31	147	2.8	4121	9 US-09-954-531-131	Sequence 131, App
32	147	2.8	4121	9 US-09-954-531-352	Sequence 352, App
33	145	2.8	4225	9 US-10-085-108-9	Sequence 9, Appl
34	145	2.8	4265	10 US-09-899-651-1	Sequence 1, Appl
35	142.5	2.7	5336	10 US-09-728-952-70	Sequence 70, Appl
36	141	2.7	3774	9 US-10-109-324-1	Sequence 1, Appl
37	139	2.7	15231	10 US-09-917-800A-1505	Sequence 1505, Ap
38	138.5	2.6	35100	10 US-09-782-378A-26	Sequence 26, Appl
39	138	2.6	7995	10 US-08-864-664-333	Sequence 333, App
40	137	2.6	4635	10 US-09-927-112-1	Sequence 1, Appl
41	136	2.6	5023	10 US-09-866-562-14	Sequence 14, Appl
42	135	2.6	8078	10 US-09-791-942-3	Sequence 3, Appl
43	134	2.6	3773	10 US-09-925-302-47	Sequence 47, Appl
44	134	2.6	5466	10 US-09-728-952-69	Sequence 69, Appl
45	134	2.6	7819	10 US-09-964-824A-239	Sequence 239, App

ALIGNMENTS

RESULT 1
US-10-033-245-21
; Sequence 21, Application US/10033245
; Patent No. US20020160392A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Godard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2930R1C7
; CURRENT APPLICATION NUMBER: US/10/033,245
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145

[illegible]

QY	81	GlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerIysLeu	100
Db	286	CAGACAGCTGCTGCTTCTGTCGACGTGTTGGCATCCCTGCTCCAGCATGACAACTC	345
QY	101	LeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGlnInAsnIleMet	120
Db	346	CTCCAGTATCCCGACACAGGACAGTGGCCACAGACCCCAACTCTGGACGACAAATCATGT	405
QY	121	AspIysAsnIyrMetAlaHisLeuValGluValGlnHisGlnArgGlyIleSerGlyIle	140
Db	406	GACAAAGATTACATGAGGCCACTGTGGAGAGTCCACATGAGAGCGGCGCCCTCCGAGAGC	465
QY	141	GlnThrPheHisSerLeuLeuThrAlaSerLeuProProArgIAspSerThrGlnAla	160
Db	466	CAGACTTTCACCTCTTGTCACAGCCTCCCTCCGCCCGCGAGACAGCACAGAGCA	525
QY	161	ProIysProIysSerSerProGluGlnProIleGlyGlnIArgIleArgIleIyrThr	180
Db	526	CCCAAAACCAAGACACACCCAGACAGCCCATGAGGCGAGGCGGATTTGGTGGGAGCC	585
QY	181	GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro	200
Db	586	CAGCTCCGGGGTGTGGGCCCTGAGAGACACTGGTGGCATGTCTCCATGATTTTCCGG	645
QY	201	LeuSerProAspProArgTyrPglInSerSerProArgProValAlaLeuAlaLeuGln	220
Db	646	CTCAGCCCGGACCTCGGTGGGACAGCTCCAGTCCCGCCCGGTGGCCCTCCCTCGAG	705
QY	221	GlnAlaLeuGlyGlnGluLeuAlaArgValValGlnIleSerProIleValProGlyIle	240
Db	706	CAGGCGCTGGGCGAGAGAGCTGGCCGGGTGTGTCCAGGACAGCCCGAGGTGCCGGGATC	765
QY	241	ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyIleAlaLeu	260
Db	766	ACGGTGGGTGTCTCGCAGGCCCTCTGCACCTCTCAGTCTCCACACAGCGGCTCCGTG	825
QY	261	ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnIleuLeuGln	280
Db	826	GTCAGTTCATGACACCTGATCCACTTCTGTGGCTGCCCTGTGTGGCCAGCTCTGGCAG	885
QY	281	TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuIysValLeuLeu	300
Db	886	TACACAGGCTGTGTGGCCACAGAGACACGGGTTCTCTGCTCTTCTTAAGGTGCTCCG	945
QY	301	GlnMetLeuGlnTrpLeuAspSerProGlyValGluGlyIleProLeuArgAlaGlnLeu	320
Db	946	CAGATGTCTGCGAGTGGCGACACCTCGGGCTGGAGGGCGGGCCCTCGGGGACACAGTTC	1005
QY	321	ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyIleLeu	340
Db	1006	AGGATGCTTGGCAGCCAGGCTCAGCCGGGCGGAGGCTCAGTGTATGTGGAGGGGGCTC	1065
QY	341	LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGluValIAspSerThrVal	360
Db	1066	CTCGCGCTGGGCGAGGCCCTCGGCTTCCGACAGAGCTGAGAGGTGTACGTCACAGCTC	1125
QY	361	ArgAlaValIleAlaThrLeuArgSerGlyGluGlnCysSerValGluProAspLeuIle	380
Db	1126	CGTCCCGTCATCGCCACCTGAGGTCTGGGGACATGACAGCTGAGCGGAGCCCTGATC	1185
QY	381	SerIysValLeuGlnGlyLeuIleGlyValArgSerProHisLeuGlnIleuLeuThr	400
Db	1186	AGCAAAAGTCTCCAGGGGGCTGATCGAGTGAAGTCCCCCACCCTGGAGAGAGCTGTGACT	1245
QY	401	AlaPhePheSerAlaThrAlaAspAlaAlaSerProPheProAlaCysIysProValVal	420
Db	1246	GCATTCTTCTCTGACATGTGGATGTGCTCCCTCCCTTTCACAGCTGTAAGCCCTGTGTG	1305
QY	421	ValValSerSerLeuLeuLeuGlnGluGluProLeuAlaGlyIleIysProIleVal	440
Db	1306	GTGTGTGAGCTTCCGTGCTGCGACGAGAGAGACCCCTGTGGTGGGGGAAAGCCGGGTGCG	1365

QY 441 AspGlySerLeuAlaValArgLeuGlyProSerSerGlyLeuLeuValAsp 460
 Db 1366 GAGGTGGACAGCTGGAGCGCGTGGCGGCCCTGTGAGGCTCTCTATGACAG 1425
 QY 461 LeuGluMetLeuAspProGluValValSerSerCysProAspLeuGluLeu 480
 Db 1426 CTGGAAGCTGGACCCGAGGTGGTCAAGACTCCCGACCTCCAGCTCAGCTGCTC 1485
 QY 481 PheSerArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 500
 Db 1486 TTCTCCCGAG 1545
 QY 501 ThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeu 520
 Db 1546 ACCCTCTCAGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1605
 QY 521 GlySerArgGluGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAla 540
 Db 1606 GGCAGAGACCGGGAAACAGAGGTTCAGACCCCTCTGAGCTCTCTGAGCTCTG 1665
 QY 541 IleHisValProArgIleTropGlnGlyArgAspGlnArgTrpProGlnLysArgGlu 560
 Db 1666 ATCCATGTCTCTGATGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1725
 QY 561 GluLeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGluLeuAla 580
 Db 1726 GAGCTGTGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 1785
 QY 581 GluAlaGluTrpArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAla 600
 Db 1786 GAGGGGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1845
 QY 601 LeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValThrGlu 620
 Db 1846 CTGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1905
 QY 621 HisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgArgCysArg 640
 Db 1906 CACCTGTGAGCTGCATCAGAGAGTGGGAGAGACAGCTGCTGGAGGGGCTGCCGAG 1965
 QY 641 LeuLeuLeuGlnLeuTrpLeuGlnArgProGluLeuArgValProValProGluVal 660
 Db 1966 CTCTCTGAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2025
 QY 661 LeuHisSerGlnGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuLeuHis 680
 Db 2026 CTGCACACGAGAGGGCTGCCAGCAGCAGCAGCTCTGCAAGCTGAGAGCTCATCACC 2085
 QY 681 PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGluAsnArgGlyAlaAsp 700
 Db 2086 TTCAATCAGCTCTCTGCGGAGACACAGCACTCCCGGCTTGGAGAACCGAGCGCGAT 2145
 QY 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuArgHisLeu 720
 Db 2146 GCCACCATGGCTGCCGGAAGCTGGGGTGGGCGCACCGCTGCTGCTGCTGCTGCTG 2205
 QY 721 ProMetIleAlaAlaLeuLeuHisGlyArgThrHisLeuAspPheGlnGluPheArg 740
 Db 2206 CCCATGATCGGGCGCTCTGACGCGCGCGACCCACCTCACTTCAGAGAGTCCGGCAG 2265
 QY 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGlnProHis 760
 Db 2266 CAGAACCCACCTGAGCTGCTCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2325
 QY 761 ValPheArgSerGluHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeu 780
 Db 2326 GTGTCCCGACGAGACACGAGGGGCGTGGGAGCTGCTCTCTCTCTCTCTCTCTCTG 2385
 QY 781 LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln 800
 Db 2386 CTGCTGAATTCAGAGAACTCTCCCGCATCTGCTGCTTATATACACAGATTGTGCGAG 2445
 QY 801 PheIleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHis 820

Db 2446 TTCATCATTAAGTACATTACATGACCCAGCAGCATCTCTCTCTGACAGAGC 2505
 QY 821 AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu 840
 Db 2506 GCCGACCCGCTCCAGACCTGCTTCCTGCAACACAGTGCATGCTGTGAAATCCCTC 2565
 QY 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgThrAspArgGlyLeuAspGlu 860
 Db 2566 CTTCGAGGGCTACGCTGCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2625
 QY 861 GlyGluGluGluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrPro 880
 Db 2626 GCGGAGGAGAGAGCTACACCGCGCTCTGCGCCGTGTCAGGCTCTCCCTGTTACACCT 2685
 QY 881 LeuThrAlaAlaGluMetAlaProTyrMetLysArgLeuSerArgGlyGlnThrValGlu 900
 Db 2686 CTGACCGGGCGGAGATGAGCCCTTACATGAAGACGCTTCCGGGGCCAAACGGTGGAG 2745
 QY 901 AspLeuLeuGluValLeuSerAspIleAspGluMetSerArgArgProGluIleLeu 920
 Db 2746 GATCTGCTGAGGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2805
 QY 921 SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGluGluCysArgAsn 940
 Db 2806 ACCTTCTCTGACCAACCTGCAGCGGCTGATGAGCTCGGCGAGAGTGTGCCGCAAC 2865
 QY 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaPhe 960
 Db 2866 CTGCGCTTCAAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2925
 QY 961 LeuProThrPheMetTyrCysLeuGlySerGlnAspPheGluValValGlnThrAlaLeu 980
 Db 2926 CTGCCACGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2985
 QY 981 ArgAsnLeuProGluTrpAlaLeuLeuCysGlnGlnHisAlaValLeuLeuHisArg 1000
 Db 2986 CGGAACCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3045
 QY 1001 AlaPheLeuValGlyMetTyrGlyGlnMetAspProSerAlaGlnIleSerGlnAlaLeu 1020
 Db 3046 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3105
 QY 1021 ArgIleLeuHisMetGlnAlaValMet 1029
 Db 3106 AGATCTCTCATATGAGGCGCTGATG 3132

RESULT 2
 US-10-033-223-21
 ; Sequence 21, Application US/10033223
 ; Patent No. US2002016464A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2930R1C9
 ; CURRENT APPLICATION NUMBER: US/10/033,223
 ; PRIOR FILING DATE: 2001-12-27
 ; PRIOR APPLICATION NUMBER: 60/095,325
 ; PRIOR FILING DATE: 1998-08-04
 ; PRIOR APPLICATION NUMBER: 60/112,851

? PRIOR FILING DATE: 1998-12-16
? PRIOR APPLICATION NUMBER: 60/113,145
? PRIOR FILING DATE: 1998-12-16
? PRIOR APPLICATION NUMBER: 60/113,511
? PRIOR FILING DATE: 1998-12-22
? PRIOR APPLICATION NUMBER: 60/115,558
? PRIOR FILING DATE: 1999-01-12
? PRIOR APPLICATION NUMBER: 60/115,565
? PRIOR FILING DATE: 1999-01-12
? PRIOR APPLICATION NUMBER: 60/115,733
? PRIOR FILING DATE: 1999-01-12
? PRIOR APPLICATION NUMBER: 60/119,341
? PRIOR FILING DATE: 1999-02-09
? PRIOR APPLICATION NUMBER: 60/119,537
? PRIOR FILING DATE: 1999-02-10
? PRIOR APPLICATION NUMBER: 60/119,965
? PRIOR FILING DATE: 1999-02-12
? PRIOR APPLICATION NUMBER: 60/162,506
? PRIOR FILING DATE: 1999-10-29
? PRIOR APPLICATION NUMBER: 60/170,262
? PRIOR FILING DATE: 1999-12-09
? PRIOR APPLICATION NUMBER: 60/187,202
? PRIOR FILING DATE: 2000-03-03
? PRIOR APPLICATION NUMBER: PCT/US99/12252
? PRIOR FILING DATE: 1999-06-02
? PRIOR APPLICATION NUMBER: PCT/US99/28634
? PRIOR FILING DATE: 1999-12-01
? PRIOR APPLICATION NUMBER: PCT/US99/28551
? PRIOR FILING DATE: 1999-12-02
? PRIOR APPLICATION NUMBER: PCT/US00/03565
? PRIOR FILING DATE: 2000-02-11
? PRIOR APPLICATION NUMBER: PCT/US00/04414
? PRIOR FILING DATE: 2000-02-22
? PRIOR APPLICATION NUMBER: PCT/US00/05841
? PRIOR FILING DATE: 2000-03-02
? PRIOR APPLICATION NUMBER: PCT/US00/08439
? PRIOR FILING DATE: 2000-03-30
? PRIOR APPLICATION NUMBER: PCT/US00/14941
? PRIOR FILING DATE: 2000-05-30
? PRIOR APPLICATION NUMBER: PCT/US00/15264
? PRIOR FILING DATE: 2000-06-02
? PRIOR APPLICATION NUMBER: PCT/US00/32678
? PRIOR FILING DATE: 2000-12-01
? NUMBER OF SEQ ID NOS: 38
? SEQ ID NO 21
? LENGTH: 3437
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-033-223-21

Alignment Scores:
Pred. No.: 0 Length: 3437
Score: 5241.00 Matches: 1029
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-929-769-7 (1-1029) x US-10-033-223-21 (1-3437)

QY 1 MeHsiIleuValaIhIsAlaMeValIleleuLeuThrlleuGluProProArGAla 20
DB 46 AlTCACATCTCTGTGGCCATGTCATGTATCTGTGCGTGGCCCGCTCGAGCC 105
QY 21 AspAspSerGluPheGlnAlaLeuLeuAspIleTrpPheProGluGluysProLeuPro 40
DB 106 GAGGACAGGAGTTCAGGCGCTGTCGACATCTGTGTTCCGAGAGAGCCACTGCC 165
QY 41 ThrlAlaPheLeuValAspThrSerGluGluAlaLeuLeuProAspTrpLeuysLeu 60
DB 166 ACCGCTTCTCTGTGACACATCGAGAGGCGCTGCTCTGCTGAGCTGAGAGCTG 225
QY 61 ArgMetIleArgSerGluValLeuArgLeuValAspAlaIleLeuGlnAspLeuGluPro 80

DB 226 CGCATGATCCGTTCTGAGGTCTCCGCTGTGAGACCCCTCGACAGACCTGAGACC 285
QY 81 GlnGlnleuLeuPheValGlnSerPheGlyIleProValSerSerMetSerLysLeu 100
DB 286 CAGCAGCTGCTGCTGTTCTGTCAGTGTGTTGGATCCCTGCTGCAGATGACAAACTC 345
QY 101 LeuGlnPheLeuAspGlnAlaValAlaIleAspProGlnThrlleuGlnIleMet 120
DB 346 CTCAGATTCTGTGACCCAGGAGTGGCCACAGCCCGACACTGTGAGAGAACATCATG 405
QY 121 AspLysAsnTrpMetAlaIleValGlnValGlnIleHisLysArgGluAlaSerGly 140
DB 406 GACAGAATTCATGAGCCACCTGTGTGAGATGACATGAGCGGCGCTCCGAGGC 465
QY 141 GlnThrPheHisSerLeuLeuThrlAlaSerLeuProProArgArgAspSerThrlAla 160
DB 466 CAGACTTTCACACTCTTGTCTACAGCTCCCTGCGCCCGCCGAGACACAGAGGCA 525
QY 161 ProLysProLysSerSerProGluGlnProIleGlyGlnGlyArgValGlyThr 180
DB 526 CCCAAACCAAGAGACAGCCAGACCCATGAGCCAGGCGGAGTTCCGCTGGGAGC 585
QY 181 GlnLeuArgValLeuGlyProGluAspAspLeuAlaIleMetPheLeuGlnIlePhePro 200
DB 586 CAGCTCGGGTGTGGCCCTGAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
QY 201 LeuSerProAspProArgTrpGlnSerSerSerProArgProValAlaLeuAlaLeuGln 220
DB 646 CTCAGCCCGGAGCCCTGCTGTCAGAGCTCCAGTCCCGCCCGCTGCGCTGCTGAG 705
QY 221 GlnAlaLeuGlnGlnLeuLeuAlaArgValValGlnGlySerProGluValProGlyIle 240
DB 706 CAGGCCCTGGCCAGAGAGTGGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
QY 241 ThrlValArgValLeuGlnAlaLeuAlaThrlleuLeuSerSerProHisGlyAlaLeu 260
DB 766 ACGGTGCTCTGTGAGGCTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 261 ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGln 280
DB 826 GTATGTCCATGCACGCTGACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuysValLeuLeu 300
DB 886 TACAGGCTGTGTGTCACAGGACACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
QY 301 GlnMetLeuGlnTrpPheLeuAspSerProGlyValGlnGlyProLeuArgAlaGlnLeu 320
DB 946 CAGATGCTGACAGTGGCTGGACAGCCCTGCTGAGGAGGCGGCGCTGCGGACACTC 1005
QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeu 340
DB 1006 AGGATGCTTCCAGCCAGGCTGAGGCTGAGGCTGAGGCTGAGTATGTCAGAGGGGCTC 1065
QY 341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGluValValSerSerThrlVal 360
DB 1066 CTGCGCTGTGCGAGGCGCTGCGCTTCTGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1125
QY 361 ArgAlaValIleAlaThrlleuAspArgSerGlyGlnGlnCysSerValGluProAspLeuIle 380
DB 1126 CGTGCCCTCATCCGACCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGATC 1185
QY 381 SerLysValLeuGlnGlnLeuIleGlnValArgSerProHisLeuGlnGlnLeuLeuThr 400
DB 1186 AGCAAGTCTCTCAGGAGGCTGATGAGTGAAGTCCCGCCAGCTGAGAGTGTGACT 1245
QY 401 AlaPhePheSerAlaThrlAlaAspAlaIleSerProPheProAlaCysLysProValVal 420
DB 1246 GCATTTCTTCTCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1305
QY 421 ValValSerSerLeuLeuGlnGlnGlnGlnProLeuAlaGlyGlyLysProGlyAla 440

Dh	1306	GGGTGAGCTCCCTGGCTGGCGACGAGAGAGAGAGCCCTTGCTGGGGGAGACCGGGGTGCG	1305
Qy	441	AspGlySerLeuGlnAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrp	460
Dh	1366	GAGGTTGGACGCTGGAGGCGCGTGGGGCCCTGTCAGGCGCTCTGATGGATCG	1425
Qy	461	LeuGlnMetLeuAspProGlnValValSerSerCysProAspLeuGlnLeuArgLeuLeu	480
Dh	1426	CTGGAAATGGCTGAGACCCCGAGGGTGGTCACAGCTGGCCCGACCTCACTCAAGCTGGCTC	1485
Qy	481	PheSerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTrpLeuLeu	500
Dh	1486	TTCTCCCGGAGAGAGGCAAGGTCAAGGCCACAGTGTCCCTGCTCCCTCACTCTG	1545
Qy	501	ThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeuLeu	520
Dh	1546	ACCCCTTTCACGCAATCAGTCCAGCTGAGCCCACTGCACACATGCATGCAGATCCTGTG	1605
Qy	521	GlyLysSerArgGlnGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys	540
Dh	1606	GGCAAGACCGGGAACAGAGGTTGCAGCCCTGCTCTGGACTTCTCTGGGGCTGC	1665
Qy	541	IleHisValProArgIleTrpGlnGlnArgAspGlnArgTrpProGlnLysArgArgGln	560
Dh	1666	ATCCATGTCTCTGCACTGTGGACGGGGCCGGACACGACGCCACCCCGACAGACGGCGAG	1725
Qy	561	GluLeuValLeuArgValGlnGlyProGlnLeuIleSerLeuValGlnLeuIleLeuAla	580
Dh	1726	GAGCTGTGCTGGGGTCCAGGGCCCGGAGTCATCAGCTGTGTGAGGTGATCTCTGGCC	1785
Qy	581	GluAlaGluThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg	600
Dh	1786	GAGCGGAGACGGGAGCCAGGACGGGGACACAGCCGCTGACGCTCATCCAGGCCGG	1845
Qy	601	LeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValThrGln	620
Dh	1846	CTGCGCTGCTGCTCACTCTCTGTGGGAGCATGATGATGATGATGATGATGATGATG	1905
Qy	621	HisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgArgCysArgAsp	640
Dh	1906	CACCTGTCAGAGCTGCATCCAGACAGTGGGGAGACAGCGTCTGGAGAGGGCTGCCGAGAC	1965
Qy	641	LeuLeuLeuGlnLeuArgTrpLeuGlnArgProGlnLeuArgValProValProGlnValLeu	660
Dh	1966	CTTCTCTGAGCTCACTACACAGCGCGGAGACTCGGGGTGCCGTGCTGATGAGTCTCA	2025
Qy	661	LeuHisSerGlnGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuLeuHisArg	680
Dh	2026	CTGCACACGAGGGGCTGCCACAGACAGCGTGTGCAACTGAGACGATCATCCACGCG	2085
Qy	681	PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnLysArgIleLysAlaAsp	700
Dh	2086	TTTCATCACGCGCTCTGGCGACACACAGACTCCCGGGGCTTGGAGACGAGGGCGGAT	2145
Qy	701	AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuLeuArgHisLeu	720
Dh	2146	GCCAGCATGGCTGCGCGAGAGCTGGCGGTGGCCACCCGCTGCTGCTCAGCACCTG	2205
Qy	721	ProMetIleAlaAlaLeuLeuHisGlyArgTrpHisLeuAspPheGlnGlnPheArgGln	740
Dh	2206	CCCATGATCGGGGCTCCGACAGGCGCGACCCACCTCAACTCCAGAGATGTCGGGAG	2265
Qy	741	GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGlnProHis	760
Dh	2266	CAGAACCACTGAGCTGCTCTTCGACAGTGTGTGGCTGCTGAGACTGTGTGACGCCGAC	2325
Qy	761	ValPheArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeu	780
Dh	2326	GTTGTTCCGACAGCGACACGAGGGGGCGCTGTGGACTGTGCTTCTGCTTCATCCGCTG	2385
Qy	781	LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln	800
Dh	2386	CTGTGTGAATTACAGGAAGTCTCCCGCATCTGTGGCTGCTTATACACAAGATTGTGCG	2445

OY	801	PhellHstLysrSYrrLIethrrYrAsnAlaIproAlaIalIeserPheLeuInlySHS	820
Db	2446	TTCATCCATAAGTACATTACCCTCAATAGCCCCAGCAGCATTCCTTGCGAAGAAGC	2505
OY	821	ALAAspProLeuHIIsApIleSerPheAspAsnSerAspLeuValMetLeuLysSerLeu	840
Db	2506	GCCGACCCGCCTCCACGACCTGTCTTGANACAAGTAGACTGGTGTATGCTGAATCCCTC	2565
OY	841	LeuAlaGlyLeuSerLeuProSerArgAspAspArgThrAspArgLyLeuAspGlu	860
Db	2566	CTTGACAGGGCTCACACCTGCCACAGCAGCAGACGACGACCGAGCGCTGGACGAAGAG	2625
OY	881	LeuThrAlaAlaGluMetAlaProTyMeLyArgLeuSerArgLyGlnThrValGlu	900
Db	2686	CTGACCGGGGGCGAGATGGCCCTTACATGAACGGGCTTCCGGGGCCAAACGGGTGAG	2744
OY	901	AspLeuLeuGluValIleuSerAspIleAspGluMetSerArgArgArgProGluIleu	920
Db	2746	GATCTGCTGGAGGTTCTGAGTGCATGAGCAGAGATGCCCGGGGAGACCCAGATCTCG	2805
OY	921	SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGluLucScyArasn	940
Db	2806	AGCTTCTTCTGACCAACCTGCACGGGCTGATGAGCTCGGCCAGAGAGTGTGGCCGCAAC	2865
OY	941	LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaIalaPhe	960
Db	2866	CTCGCCTTACACCTGGCCCTGGCTCCATGCAGAACGCCCGCAGCATTCGACGGCTTTC	2922
OY	961	LeuProthrPheMetTYrCySLeuGlySerGlnAspPheGluValGlnThrAlaLeu	980
Db	2926	CTGCCACGTTCACTACGTGCTGGCGAGCAGACACTTGTAGGTGGTGCAGACGGCCCTC	2983
OY	981	ArgAsnLeuProGluTYrAlaLeuLeuCysGlnGluHisAlaIalValLeuLeuHisArg	1000
Db	2986	CGGAACTCGCTGAGTACGCTCTCTGTCGMAAGACGCGGGCTGTGCTGCCACCGG	3044
OY	1001	AlaPheLeuValGlyMetTYrGlyGlnMetAspProSerAlaGlnIleserGluAlaLeu	1020
Db	3046	GCCTTCCGTGGGGCATGTACGGCCAGATGAGACCCGACCGCATCTCCGAGGGCCCTG	3105
OY	1021	ArgIleLeuHisMetGluAlaValMet	1029
Db	3106	AGGATCTGCATATGAGGCGCGTGATG	3132
RESULT 3			
US-10-033-167-21			
Sequence 21, Application US/10033167			
Publication No. US20020182618A1			
GENERAL INFORMATION:			
APPLICANT: Botstein, David			
APPLICANT: Desnoyers, Luc			
APPLICANT: Ferrara, Napoleone			
APPLICANT: Fong, Sherman			
APPLICANT: Gao, Wei-Qiang			
APPLICANT: Goddard, Audrey			
APPLICANT: Gurney, Austin L.			
APPLICANT: Pan, James			
APPLICANT: Roy, Margaret Ann			
APPLICANT: Stewart, Timothy A.			
APPLICANT: Tumas, Daniel			
APPLICANT: Watanabe, Collin K.			
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic			
TITLE OF INVENTION: Acids Encoding the Same			
FILE REFERENCE: P2930R1C10			
CURRENT APPLICATION NUMBER: US/10/033,167			
CURRICULUM FILING DATE: 2001-12-27			
PRIOR APPLICATION NUMBER: 60/095,325			

```

: PRIOR FILING DATE: 1998-08-04
: PRIOR APPLICATION NUMBER: 60/112,851
: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,145
: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,511
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: 60/115,558
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/115,565
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/115,733
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/119,341
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: 60/119,537
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: 60/119,965
: PRIOR FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: 60/162,506
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: 60/170,262
: PRIOR FILING DATE: 1999-12-09
: PRIOR APPLICATION NUMBER: 60/187,202
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: PCT/US99/28634
: PRIOR FILING DATE: 1999-12-01
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US00/03565
: PRIOR FILING DATE: 2000-02-11
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: 2000-03-02
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: PCT/US00/14941
: PRIOR FILING DATE: 2000-05-30
: PRIOR APPLICATION NUMBER: PCT/US00/15264
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: 2000-12-01
: NUMBER OF SEQ ID NOS: 38
: SEQ ID NO 21
: LENGTH: 3437
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-033-167-21

Alignment Scores:
Pred. No.: 0          Length: 3437
Score: 5241.00       Matches: 1029
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%    Indels: 0
DB: 9                Gaps: 0

US-09-929-769-7 (1-1029) x US-10-033-167-21 (1-3437)
QY 1 MethistilleuVal1H1sAlaMetVal1LeuLeuThrlLeuGlyProProAla 20
Db 46 ATGCACATCTCTCGTGGTCCATGCGATGATCTCTGAGAGCGTGGCCGCTCGAGCC 105
QY 21 AspAspserGluPheGlnAlaLeuAsp1LeTrpPheProGluGluValysProLeu 40
Db 106 GACACACGCGAGTTCAGGCGCTGCTGACATCTGTTCCGGAGGAGAACCCACTGCC 165
QY 41 ThrAlaPheLeuValAspThrSerGluGluAlaLeuLeuProAspTrpLeuLysLeu 60
Db 166 ACGCCTTCTCGTGGAGACATGAGAGGCGCTGCTGCTCTGACTGACTGACTG 225

```

```

QY 61 ArgMetIleArgSerGluValLeuArgLeuValAspAlaAlaLeuGlnAspLeuPro 80
Db 226 CGCATGATTCGTTCTGAGGTCTCTCGCTGCTGAGCGCGCCGTCGAGAGACTGAGACCG 285
QY 81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerLysLeu 100
Db 286 CAGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
QY 101 LeuGlnPheLeuAspGlnAlaValAlaH1sAspProGlnTrpLeuGluGlnAlaMet 120
Db 346 CTCAGTTCCTGAGACAGGAGTGGCCAGACCCGACAGCTGAGACAGACATCATG 405
QY 121 AspLysAsnTyrMetAlaH1sLeuValGluValGlnH1sGluArgGlyAlaSerGlyGly 140
Db 406 GACAAGAATTCATGAGGCCACCTGCTGAGAGTCCAGCATGAGCGCGCGCTCCGAGGC 465
QY 141 GlnThrPheH1sSerLeuLeuThrlAlaSerLeuProProAlaArgAspSerThrlAla 160
Db 466 CAGACTTCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
QY 161 ProLysProLysSerSerProGluGlnProIleGlyGlnArgValGlyValGly 180
Db 526 CCAAAACCAAGACAGACCCGAGACAGCCATAGGCCAGGCGCGGATTCGGGTGGGACC 585
QY 181 GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro 200
Db 586 CAGCTCGGCGTGTGCGGCGCTGAGAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
QY 201 LeuSerProAspProAlaGlyTrpGlnSerSerSerProAlaAlaLeuGln 220
Db 646 CTCAGCCGCGACCTCGTGGTGGAGAGCTCCAGTCCCGCGCGCGCTGCGCTGCTGCTGCTG 705
QY 221 GlnAlaLeuGlyGlnGluLeuAlaArgValAlaGlnLysSerProGluValProGlyIle 240
Db 706 CAGGCGCTGGGCGCAGAGAGTGGCGCGGTGCTCCAGGCGACCCGAGAGTGGCGGCGATC 765
QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProH1sGlyValAlaLeu 260
Db 766 ACGGTGGTGTCTGCGAGCGCTGCGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 261 ValMetSerMetH1sArgSerH1sPheLeuAlaLysProLeuLeuArgGlnLysGln 280
Db 826 GTGATGTCATGACACCTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeuLeu 300
Db 886 TACCAGGCGTGTGTGCGACAGACAGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
QY 301 GlnMetLeuGlnTrpLeuAspSerProGlyValGlnGlyGlyProLeuArgAlaGlnLeu 320
Db 946 CAGATGCTGACGTGGTGGAGACCGCTGCGTGGAGGCGGCGCGCTGCGGCGACACTC 1005
QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyIleLeu 340
Db 1006 AGGATGCTTGGACACAGCGCTCAGCCGCGGCGAGGCTCACTGATGTGCCAGGGGGGCTC 1065
QY 341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGluValSerSerThrVal 360
Db 1066 CTGCGCGTGGCGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
QY 361 ArgAlaValIleAlaThrLeuArgSerGlyGlnLysSerValGluProAspLeuIle 380
Db 1126 CGTCCGCTGATCCAGCGACCTGAGGTGTGGGAGAGAGTGCACCGTGGAGCGGACTATC 1185
QY 381 SerLysValLeuGlnGlyLeuIleGluValArgSerProH1sLeuGluGluLeuLeuThr 400
Db 1186 AGCAAACTCTCCAGGCGCTGATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 1245
QY 401 AlaPhePheSerAlaThrAlaAspAlaLaserProPheProAlaCysLysProValVal 420
Db 1246 GCATTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1305

```

[illegible]

Db	2386	CTGCTGATTTACAGGAAGTCCTCCGCCATCTTGCGCTGCCCTTCATCAACAATTTGTGCAG	2445
QY	801	PheIleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHis	820
Db	2446	TTTCATCCATTAAGTACATTACTACATATGCCCCAGCAGCAATCTCTTCTCACAAGACAC	2505
QY	821	AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu	840
Db	2506	GCCACACCGGCTCCACAGCAGCTGCTCTTCGACCAACAGTGAAGCTGTGTATCTGAATTCCTC	2565
QY	841	LeuAlaGlyLeuSerLeuProSerArgAspAspArgTrpAspArgGlyLeuAspGlu	860
Db	2566	CTTCAGAGGCTCAGCCTGCCACAGGAGCGACAGGACCGAGCCGAGGCTTGACCAAGAG	2625
QY	861	GlyGlnGlnLysSerSerAlaGlySerLeuProLeuValSerValSerLeuPheTrpPro	880
Db	2626	GCCGAGAGAGAGAGCTCAGCCGCCGCTCTTGCCCTGTGACGCTGCTCCCTTCACCCCT	2685
QY	881	LeuThrAlaAlaGluMetAlaProTyrMetLysArgLeuSerArgGlyGlnThrValGlu	900
Db	2686	CTGACCCCGGCGCAGAGATGGCCCCCTTCACATGAACAGCGTTTCCCGGGCCAAACGCTGGAG	2745
QY	901	AspLeuLeuGlnValLeuSerAspIleAspGluMetSerArgArgArgProGluIleLeu	920
Db	2746	GATCTGCTGGAGGTTCTGAGTGAACATAGACAGAGATGTCCCGCGGAGACCCGAGATCCTG	2805
QY	921	SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGluGlnLysCysArgAsn	940
Db	2806	AGCTTCTTTCGACCAACCTGCAGCGGGGTGATGAGTGGCGGAGAGAGTGTGCCGCGAC	2865
QY	941	LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaPhe	960
Db	2866	CTCCCTTCACAGCCTGGCGGCTCCGCTCCATGACGAACAGCCCGACGATTCGACGCGCTTC	2925
QY	961	LeuProThrPheMetTrpCysLeuGlnLysArgIleAspPheGluValGlnThrAlaLeu	980
Db	2926	CTGCGCCACGTTCAATGACTCTGCGGGCAGCCAGGACTTGAGAGTGTGAACAGCGCCTTC	2985
QY	981	ArgAsnLeuProGluTrpAlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArg	1000
Db	2986	CGGAACCTGCGCTAGTACGCTCCCTCGTGCACAAAGACACGCGGCTGTCTCTCCACCG	3045
QY	1001	AlaPheLeuValGlyMetTrpGlyGlnMetAspProSerAlaGlnIleSerGluAlaLeu	1020
Db	3046	GCCCTCTCGTGGGCAATGTACGCGCCAGATGGACCCACGCGCAGATCTCCGAGGCGCTG	3105
QY	1021	ArgIleLeuHisMetGluAlaValMet	1029
Db	3106	AGGATCCTGCATATGAGGCGCGTGAATG	3132
RESULT 4			
US-10-033-244-21			
Sequence 21, Application US/10033244			
Publication No. US2002019268A1			
GENERAL INFORMATION:			
: APPLICANT: Botstein, David			
: APPLICANT: Desnoyers, Luc			
: APPLICANT: Ferrara, Napoleone			
: APPLICANT: Fong, Sherman			
: APPLICANT: Gao, Wei-Qiang			
: APPLICANT: Goddard, Audrey			
: APPLICANT: Gurney, Austin L.			
: APPLICANT: Pan, James			
: APPLICANT: Roy, Margaret Ann			
: APPLICANT: Stewart, Timothy A.			
: APPLICANT: Tumas, Daniel			
: APPLICANT: Watanabe, Colin K.			
: APPLICANT: Wood, William I.			
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic			
TITLE OF INVENTION: Acids Encoding the Same			
FILE REFERENCE: P2930R1C2			
CURRENT APPLICATION NUMBER: US/10/033,244			

;; CURRENT FILING DATE: 2001-12-27
;; PRIOR APPLICATION NUMBER: 60/095,325
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/112,851
;; PRIOR FILING DATE: 1998-12-16
;; PRIOR APPLICATION NUMBER: 60/113,145
;; PRIOR FILING DATE: 1998-12-16
;; PRIOR APPLICATION NUMBER: 60/113,511
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/115,558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115,565
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115,733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/119,341
;; PRIOR FILING DATE: 1999-02-09
;; PRIOR APPLICATION NUMBER: 60/119,537
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/119,965
;; PRIOR FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: 60/162,506
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: 60/170,262
;; PRIOR FILING DATE: 1999-12-09
;; PRIOR APPLICATION NUMBER: 60/187,202
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: 1999-06-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28634
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/28551
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: 2000-02-11
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: 2000-02-22
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: 2000-03-30
;; PRIOR APPLICATION NUMBER: PCT/US00/14941
;; PRIOR FILING DATE: 2000-05-30
;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; NUMBER OF SEQ ID NOS: 38
;; SEQ ID NO 21
;; LENGTH: 3437
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-033-244-21

Alignment Scores:
Pred. No.: 0 Length: 3437
Score: 5241.00 Matches: 1029
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-929-769-7 (1-1029) x US-10-033-244-21 (1-3437)

QY 1 MethisIleuValValHisAlaMetValIleLeuLeuThrLeuGlyProPArgAla 20
DB 46 ATGCATCATCTGCTGTCATGCCATGTCCTGCTGACGCTGGGCGCCGCTCGAGCC 105

QY 21 AspAspSerGluPheGlnAlaLeuLeuAspIleTirPheProGluGlyProLeuPro 40
DB 106 GACGACAGGAGATTCCAGGGGCTGTCGACATCTGTTCCGAGAGAGACCATGCGCC 165

QY 41 ThrAlaPheLeuValAspThrSerGluGluAlaLeuLeuProAspThrLeuLysLeu 60

DB 166 ACCGCTTCCCTGGTGGACATCGAGAGAGCCCTGCTGCTCTGCTGACGTCGAGAGCTG 225
QY 61 ArgMetIleArgSerGluValIleArgLeuValAspAlaIleGlnAspLeuGluPro 80
DB 226 GCGATGATCCGTTCTGTGAGGTCTCCGCTGTGTGACACCCCTGCGAGGACGTGGAGCCG 285
QY 81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyTLeProValSerSerMetSerLysLeu 100
DB 286 CAGACAGCTGCTGCTGTCGTCGAGGTGTTGGCATCCCGCTGTCGACATGACCAAACTC 345
QY 101 LeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGluGlnAsnIleMet 120
DB 346 CTCACGTTCTTGACACGAGCGAGTGGCCACAGCCCTGCTGAGCGAAGAACATCATG 405
QY 121 AspLysAsnTyrMetAlaHisIleValGlnHisGlnArgGlyAlaSerGlyGly 140
DB 406 GACAGAGATTACATGGCCCACTGGTGGAGTTCACAGTATAGCGCGGCTCCGAGAGC 465
QY 141 GlnThrPheHisSerLeuLeuThrAlaSerLeuProProArgArgAspSerThrGluAla 160
DB 466 CAGACTTCCACTCTGTCACAGCCTCCCTGCGCCGCGGAGACACACAGAGCA 525
QY 161 ProLysProLysSerSerProGluGlnProIleGlyGlnGlyArgTleArgValGlyThr 180
DB 526 CCCAAACCAAGAGACACCCAGAGACCCATAGGCCAGGCGGATCGGGTGGAGAC 585
QY 181 GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro 200
DB 586 CAGCTCGGGTGTGGGCGCTGAGAGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
QY 201 LeuSerProAspProArgTirGlnSerSerSerProArgProValAlaLeuAlaLeuGln 220
DB 646 CTCAGCCCGGAGCCCTGCTGTCGAGAGTCCAGTCCCGCCCTGCTGCTGCTGCTGCTGCTG 705
QY 221 GlnAlaLeuGlyGlnGluIleAlaArgValValGlnGlySerProGluValProGlyTle 240
DB 706 CAGGCTTGGGCGGAGAGCTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyAlaLeu 260
DB 766 ACGGTGCTGCTGCTGCGAGGCTCGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 261 ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGln 280
DB 826 GTATGTCATGACCGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuLysValLeuLeu 300
DB 886 TACACAGCTGTGTGCTGACAGACACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
QY 301 GlnMetLeuGlnThrLeuAspSerProGlyValGluGlyProLeuArgAlaGlnLeu 320
DB 946 CAGATGCTGACGTGCTGCGACAGACACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeu 340
DB 1006 AGGATGCTTGGCCAGCGAGGCTTCCAGCGGCGGAGCTCATGATGTGCGAGGGGGCTC 1065
QY 341 LeuArgLeuAlaGluAlaLeuAlaPheArgGlnAspLeuGluValValSerSerThrVal 360
DB 1066 CTCGCGCTGGCGAGGCGCTGCGCTTCCGTGAGGACTGTGAGGTGTGCTGCTGCTGCTG 1125
QY 361 ArgAlaValIleAlaThrLeuArgSerGlyGluGlnCysSerValGluProAspLeuIle 380
DB 1126 CGTGGCGTATGCGCACACCTGAGGCTGCGGAGACAGTGCAGGCGGAGCGACCTGATC 1185
QY 381 SerLysValLeuGlnGlyLeuIleGluValArgSerProHisLeuGluLeuLeuThr 400
DB 1186 AGCAAGTCTCTCCAGGGGCTGATGAGGTGAGTCCCGCCACTGTGAGAGAGCTGCTGACT 1245
QY 401 AlaPhePheSerLThrLhrAlaAspAlaIleSerProPheProAlaCysLysProValVal 420


```

FILE REFERENCE: P2930R1C5
CURRENT APPLICATION NUMBER: US/10/033.435
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095.325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112.851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113.145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113.511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115.558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115.565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115.733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119.341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119.537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119.965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162.506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170.262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187.202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 21
LENGTH: 3437
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-435-21

Alignment Scores:
Pred. No.: 0
Score: 5241.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9
Gaps: 0

US-09-929-769-7 (1-1029) x US-10-033-435-21 (1-3437)
QY 1 MetHisIleuValValHisAlaMetValIleuLeuThrLeuGlyProProAla 20
DB 46 ATGCACATCTCGTGGTCCATGCAATGCAATCTCTGACGCGTGGCCCGCTCGAGCC 105
QY 21 AspAspSerGluPheGlnAlaLeuLeuAspIleTrpPheProGluGluProLeuPro 40
DB 106 GAGCAGACGCGATTCCAGGGGCTGCTGACATCTGTTCCGAGAGAGAACCCACTGCC 165

```

```

QY 41 ThrAlaPheLeuValAspThrSerGluGlnAlaLeuLeuProAspTrpLeuVal 60
DB 166 ACCGCTTCTCGTGGACACATCGAGAGGCGGCTGCTTCTCTGACTGACTGACCTG 225
QY 61 ArgMetIleArgSerGluValLeuArgValAspAlaAlaLeuGlnAspLeuPro 80
DB 226 CGCATGATCCGTTCTGAGGTCCTCCGCTGAGACCCCGCTGACAGACCTGAGACCG 285
QY 81 GlnIleuLeuLeuPheValGlnSerPheGlyIleProValSerMetSerLeu 100
DB 286 CAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
QY 101 LeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGlnIleMet 120
DB 346 CTCAGATTCTCGACACAGCAGTGGCCAGACCCCGACTGAGAGAACATCATG 405
QY 121 AspLysAsnTrpMetAlaHisLeuValGluValGlnHisGluArgGlyAlaSerGly 140
DB 406 GACAGAGATTACATGATGCCACCTGCTGAGATCCAGATGAGCCGCGCTCCGAGGC 465
QY 141 GlnThrPheHisSerLeuLeuThrAlaSerLeuProProArgAspSerThrGluAla 160
DB 466 CAGACTTCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
QY 161 ProLysProLysSerSerProGluGlnProIleGlyGlnGlyArgValGlyThr 180
DB 526 CCCAAGCCAAAGACACACCCAGACACCCATAGCCAGGCGCGGATTCGGGTGGGAC 585
QY 181 GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheGlnIlePhePro 200
DB 586 CAGCTCCGGGTGCTGGGCGCTGAGAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
QY 201 LeuSerProAspProArgTrpGlnSerSerSerProArgProValAlaLeuGln 220
DB 646 CTCAGCCCGGACCTCGTGGTGGACAGCTCCAGTCCCGCCCGGCGCTGCTGCTGCTGCTG 705
QY 221 GlnAlaLeuGlyGlnGluLeuAlaArgValAlaGlnGlySerProGluValProGly 240
DB 706 CAGGCTCGGGCCAGGAGCTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyAlaLeu 260
DB 766 AGGCTGGGTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 261 ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGln 280
DB 826 GTGATGTCATGCAACGCTAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeuLeu 300
DB 886 TACAGCGCTGTGTGCCACAGGACACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
QY 301 GlnMetLeuGlnTrpLeuAspSerProGlyValGlnGlyProLeuArgAlaGlnLeu 320
DB 946 CAGATGCTGCAATGCTGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeu 340
DB 1066 AGGATGCTTGCACACCGCTGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
QY 341 LeuArgLeuAlaGluAlaLeuAlaPheArgGlnAspLeuGluValAlaSerSerThrVal 360
DB 1066 CTGGCGCTGGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
QY 361 ArgAlaValAlaIleAlaThrLeuArgSerGlyGlnGlnCysSerValGluProAspLeuIle 380
DB 1126 CGTGGCTGATCGCACCTGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1185
QY 381 SerLysValLeuGlnGlyLeuIleGluValAlaArgSerProHisLeuGlnGluLeuThr 400
DB 1186 AGCAAACTCTCCAGAGGCGTGAATGAGTGAAGTCCCGACCTGAGAGAGAGAGTGTGACT 1245

```

QY 401 AlaPheSerAlaThrAlaAspAlaAspProPheProAlaCysLysProValVal 420
 DB 1246 GCATCTCTCTGCGACATCGGATCTGCTCCCGTTTCCAGCTGTACGCCCTGTGG 1305
 QY 421 ValValSerSerLeuLeuLeuGlnGlnGlnProLeuAlaGlyLysProGlnAla 440
 DB 1306 GTGGTGCATCTCTCTCTGTCAGAGAGAGAGCCCTGGCTGGGGAGAGCCGGTGG 1365
 QY 441 AspGlyGlySerLeuGlnAlaValAlaArgLeuGlyProSerSerGlyLeuLeuValAspTrp 460
 DB 1366 GACGGTGCAGCTCTGAGAGCCGTGGCTGGGGCCCTGTCAGGCTCTCTATGTGAGACTGG 1425
 QY 461 LeuGlnMetLeuAspProGlnValValSerSerCysProAspLeuGlnLeuArgLeuLeu 480
 DB 1426 CTGGAATAGCTGAGCCCGAGGTGGTCAAGACACTGCCGACCTGACGTCAAGCTGTCTC 1485
 QY 481 PheSerArgArgGlyGlyLysGlyGlnAlaGlnValProSerPheArgProTrpLeuLeu 500
 DB 1486 TTCTTCCCGAGAGAGAGGCAAGGTGAGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1545
 QY 501 ThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeuLeu 520
 DB 1546 ACCCTTCTACGATCAGATCCAGCTGGCCACACTGACAGTGCATCCGAGTCTGCTG 1605
 QY 521 GlyLysSerArgGlnGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys 540
 DB 1606 GGCAGAGACCGGGAGACAGAGGTGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1665
 QY 541 IleHisValProArgIleTrpGlnGlyArgAspGlnArgTrpProGlnLysArgArgGlu 560
 DB 1666 ATCCATGTCTCTGATCTGAGAGGGGGGAGCAGCCAGCCGCCAGACAGGGGGGAG 1725
 QY 561 GlnLeuValLeuArgValGlnGlyProGlnLeuIleSerLeuValGlnLeuLeuLeuAla 580
 DB 1726 GAGCTGGTGTCTGCGGCTCCAGGGCCCGGAGCTCATCAGCTGTGAGACTGATCTGCGC 1785
 QY 581 GlnAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg 600
 DB 1786 GAGCGGAGAGCGCGGAGCAGAGCGGGGACACAGCCGCTGACGCTCATCCAGGCGCGG 1845
 QY 601 LeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValThrGlu 620
 DB 1846 CTGCCCCGTGCTCTAGCTGTGCTGTGGGAGCATGAGATGTGAGAGAGGTGACGGAG 1905
 QY 621 HisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgArgCysArgAsp 640
 DB 1906 CACCTGTAGGCTGCATTCAGAGAGTGGGAGACACAGCTGTGGAGAGCGCTGCCGAGAC 1965
 QY 641 LeuLeuLeuGlnLeuTrpLeuGlnArgProGlnLeuArgValProValProGlnValLeu 660
 DB 1966 CTTCCTGTGACAGCTTACCTACAGCGCGGAGCTGGGGTGCCTGCTGCTGAGGTCTTA 2025
 QY 661 LeuHisSerGlnGlyAlaAlaSerSerValCysLysLeuAspGlyLeuLeuHisArg 680
 DB 2026 CTGCAACAGCGAGGGGTGCTCCAGCAGCAGCGCTCTCAACCTGAGAGGACTCATCACCC 2085
 QY 681 PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnAsnArgGlyAlaAsp 700
 DB 2086 TTCACTACAGCTCTTGGCGAGACACAGCACTCCCGGGCTTGGAGAGACGAGGGCGGAT 2145
 QY 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuArgHisLeu 720
 DB 2146 GCCACATAGGCTCTGCGGAGAGCTGGCGGAGCGGACCCGCTGCTGCTGAGCGACCTG 2205
 QY 721 ProMetIleAlaAlaLeuLeuHisGlyArgTrpHisLeuAsnPheGlnGlnLeuArgGln 740
 DB 2206 CCCATGATGCGGGCTCTGACGCGCGCACCCCACTCAACTTCAGAGAGTTCGGGCGAG 2265
 QY 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuLeuInProHis 760
 DB 2266 CAGAACACACCTAGAGTGTCTCTGACAGCTGTGGGCTGTGTGAGCTGTGTGACCGCGCAC 2325
 QY 761 ValPheArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeu 780

DB 2326 GTGTTCCGAGGAGACACAGGGGGGCGTGTGGAGCTGCTTCTGCTCATCCGCCCTG 2385
 QY 781 LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln 800
 DB 2386 CTGCTGAATTTACAGGAAGTCTCCCGCATCTGCGCTGCTTCATCAACAAGTTTGTGACG 2445
 QY 801 PheIleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHis 820
 DB 2446 TTCACTCAATAGTACATTACCTTACATGCCCCAGCAGCCATCTCTCTCTGAGAGAGCAC 2505
 QY 821 AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu 840
 DB 2506 GCCGACCGCTCCAGAGCTGATCTTCGACAAACAGAGACTGTGTATGCTGAATCCCTTC 2565
 QY 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgTrpHisAspArgLysLeuAspGlnGlu 860
 DB 2566 CTGGCAGGGCTCAGCTTGGCTCCAGCAGGGAGCAGACAGGACCGAGGCTGTGAGCGAAGAG 2625
 QY 861 GlyGlnGlnGlnSerSerAlaGlySerLeuProLeuValSerValSerLeuPheTrpPro 880
 DB 2626 GCGGAGAGAGAGCTCAAGCCGGCTCTTGGCCCTGTGTGACGCTCTCTGTTACACCTT 2685
 QY 881 LeuThrAlaAlaGlnMetAlaProTrpMetLysArgLeuSerArgGlyGlnThrValGlu 900
 DB 2686 CTGACCGCGCGCGAGATGGCCCCCTACATGAACAGGCTTCCCGGGGCCAACGGTGGAG 2745
 QY 901 AspLeuLeuGlnValLeuSerAspIleAspGlnMetSerSerArgArgArgProGlnIleLeu 920
 DB 2746 GATCTGTGGAGATCTGTGATGACATGAGAGATGTCCCGCGGAGACCGAGATCTTG 2805
 QY 921 SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGlnLysCysArgAsn 940
 DB 2806 ACCTTCTTCTCCAGCACTGTCAGCGGCTGATGAGTGTGCGCGAGAGAGTGTGCCGACAC 2865
 QY 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaAlaPhe 960
 DB 2866 CTGCGCTTCAAGCTGCGCCCTGCGCTCATGACAGAACAGCCAGACTTGCACCGCTTTC 2925
 QY 961 LeuProThrPheMetTrpCysLeuGlnGlySerGlnAspPheGlnValValGlnThrAlaLeu 980
 DB 2926 CTGCCCCATGTATGATGCTGCTGGGACAGCAGACTTGTGAGGTGTGAGAGGCGCTTC 2985
 QY 981 ArgAsnLeuProGlnTrpTrpAlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArg 1000
 DB 2986 CGGAACCTGCTCGAGTACCTCTCTGTGCCAAGACAGCGGCTGTGCTGCCACCGG 3045
 QY 1001 AlaPheLeuValGlyMetTrpGlyGlnMetAspProSerAlaGlnIleSerGlnAlaLeu 1020
 DB 3046 GCCTTCTGTGGGATGTGACGGCCAGATGAGACCCAGCGCGCATCTCCAGAGGCGCTG 3105
 QY 1021 ArgIleLeuHisMetGlnAlaValMet 1029
 DB 3106 AGGATCTCTCATATGAGGCGGTGATG 3132

RESULT 6
 US-10-032-990-21
 ; Sequence 21, Application US/10032990
 ; Publication No. US20030032060A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumms, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.

Db 1186 AGCAAGTCTCCAGGCGCTGATCGAGGTGAGGTCCCTCCAGCTGAGAGCTGCTGACT 1245
 QY 401 AlapheSerSerAlaThrAlaAspAlaSerProPhaAlaGlySerProValVal 420
 Db 1246 GCAATCTCTCTGCGACCTGGAGTGTCTGCTCCCGCTTTCAGCTGTAAGCCCGCTGTG 1305
 QY 421 ValValSerSerLeuLeuLeuGlnGlnGlnGlnProLeuAlaGlyLysProGlyAla 440
 Db 1306 GGGGTGAGCTCTCTGCTGTCAGAGAGAGAGAGCCCTGCTGGGGGAAAGCGGGTGGC 1365
 QY 441 AspGlyGlySerLeuGlnAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrp 460
 Db 1366 GAGGTGGAGCGCTGAGAGCGCTGCGGGCGCTGCTGAGCGCTCTAGAGGACTGG 1425
 QY 461 LeuGlnMetLeuAspProGlnValValSerSerCysProAspLeuGlnLeuArgLeuLeu 480
 Db 1426 CTGGAAATCTGACCCAGGAGGTGTCACAGCTGCCCCGACCTGAGCTCAGAGCTGCTC 1485
 QY 481 PheSerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTyrLeuLeu 500
 Db 1486 TTCTCCGGAGAGAGGCAAGAGTCAAGGCGCCAGGTCCTCGTCCGCTCCATACCTCTG 1545
 QY 501 ThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysTlaArgValLeuLeu 520
 Db 1546 ACCCTCTTACGCATCAGTCCAGCTGGCCACACTGACCAGTGCATCCGAGTCTGCTG 1605
 QY 521 GlyLysSerArgGlnGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys 540
 Db 1606 GCGAAGAGCGGGAAACAGAGTTCGACCCCTGCTCTGCTGACTTCTCTGAGGCTGCTG 1665
 QY 541 IleHisValProArgIleTrpGlnGlyArgAspGlnArgTrpProGlnHisArgArgGln 560
 Db 1666 ATCCATGTCCTCGCATCTGGCAGGGGGCGGAGCACAGCGACCCCGCAAGAGCGGGAG 1725
 QY 561 GlnLeuValLeuArgValGlnGlyProGlnLeuIleSerLeuValGlnLeuIleLeuAla 580
 Db 1726 GAGCTGTGCTCTGGGTCCAGGGCCCGGAGCTCATCAGCTGTGGAGTGTGATCCCTGG 1785
 QY 581 GlnAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg 600
 Db 1786 GAGCGCGAAGCCGGAGCGGAGAGCGGAGACAGCGCGCTGCACTCATCCAGCGCCG 1845
 QY 601 LeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValIleArgGln 620
 Db 1846 CTGCCCCCTGCTGCTGAGCTGTGCTGGGAGCATGAGTGTGAGGAGTGTGAGGAG 1905
 QY 621 HisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlnLysArgCysArgAsp 640
 Db 1906 CACCTGTCTAGGCTGCATCCAGAGTGGGAGACAGCTGTGGAGAGGCGCTGCCAGAC 1965
 QY 641 LeuLeuLeuGlnLeuTyrLeuGlnArgProGlnLeuArgValProValProGlnValLeu 660
 Db 1966 CTTCCTCTGACCTTACTTACAGCGGCGGAGCTCGGGTGGCTGCTGAGTGTCTA 2025
 QY 661 LeuHisSerGlnGlyAlaAlaSerSerSerValCysLysLeuAspLysLeuIleHisArg 680
 Db 2026 CTGCACAGCGAAGGGGCTCCAGACAGCAGCTGTCCAGAGTGTGAGACTCATCCACCG 2085
 QY 681 PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnLysArgGlyAlaAsp 700
 Db 2086 TTTCATCAGCTCTCTGCGACACAGCTCCCGGGCTGTGGAGAACCGAGGCGGGAGT 2145
 QY 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuLeuArgHisLeu 720
 Db 2146 GCCAGCATGTGCTGCGGAAAGTGGGGTGGGCGACCCGCTGCTGTCTCAGACGCTG 2205
 QY 721 PrometIleAlaAlaLeuLeuHisGlyArgTrpHisLeuAspPheGlnLysPheArgGln 740
 Db 2206 CCCATGATGCGGCGCTCTGACAGGCGCACCCACCTCAACTCCAGAGGTTCGGGAG 2265
 QY 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGlnProHis 760
 Db 2266 CAGAACCACTGAGCTGTCTGACAGCTGTGGGCTGTGAGACTGTGTGACGCGGAC 2325

QY 761 ValPheArgSerGlnHisGlnGlnValAlaLeuTrpAspCysLeuLeuSerPheIleArgLeu 780
 Db 2326 GTTCTCCAGAGCAGACACAGGGGGCGCTGTGGAGCTGCTTCTGCTTCACTCCGCTG 2385
 QY 781 LeuLeuAsnTyrArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln 800
 Db 2386 CTGCTCAATTACAGAAATGCTCCCGCATCTGCTGCTGCTATCAAAAGTTGTGACAG 2445
 QY 801 PheIleHisLysTyrIleThrTyrAsnAlaProAlaAlaIleSerPheLeuGlnLysHis 820
 Db 2446 TTTCATCATAAGTACATTAATTAATGCCCCAGCAGCCATCTCTCTCTGCGAAGGAC 2505
 QY 821 AlaAspProLeuHisAspLeuSerPheAspAspSerAspLeuValMetLeuLysSerLeu 840
 Db 2506 GCCAGCCGCTCCAGCAGCTGCTCTTGCAGACAGTGAACCTGTGATGCTGAATATCCCT 2565
 QY 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgTrpAspArgLysLeuAspGlnGln 860
 Db 2566 CTTCAGAGGCTCAGCTGCTCCAGCAGAGAGAGACAGACCGAGCGGCTGAGCAAGAG 2625
 QY 861 GlyGlnGlnGlnSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrPro 880
 Db 2626 GGGAGAGAGAGAGCTCAGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2685
 QY 881 LeuThrAlaAlaGlnMetAlaProTyrMetLysArgLeuSerArgGlyGlnThrValGln 900
 Db 2686 CTACAGCGCGCGCAGATGCGCCCTTACATGAACGCTTCCGGGGCCAAACGCTGGAG 2745
 QY 901 AspLeuLeuGlnValLeuSerAspIleAspGlnMetSerArgArgArgProGlnIleLeu 920
 Db 2746 GATCTCTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2805
 QY 921 SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerIleGlnGlyCysArgAsp 940
 Db 2806 AGCTTCTTCTCGACCACTGCGCAGCGCTGATGATGATGATGATGATGATGATGATG 2865
 QY 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaAlaPhe 960
 Db 2866 CTGCGCTTCAAGCTGCGCCCTGCTCATGACAGACAGCCGAGCATTTGAGCGCTTC 2925
 QY 961 LeuProThrPheMetTyrCysLeuGlySerGlnAspPheGlnValAlaGlnThrAlaLeu 980
 Db 2926 CTGCCCCAGCTTCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2985
 QY 981 ArgAspLeuProGlnTyrAlaLeuLeuCysGlnGlnHisAlaValLeuLeuHisArg 1000
 Db 2986 CGGAACCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3045
 QY 1001 AlaPheLeuValGlyMetTyrGlyGlnMetAspProSerAlaGlnIleSerGlnAlaLeu 1020
 Db 3046 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3105
 QY 1021 ArgIleLeuHisMetGlnAlaValMet 1029
 Db 3106 AGGATCTGCTGATGAGGCGCTGATG 3132

RESULT 7
 US-09-929-769-3
 ; Sequence 3, Application US/09929769
 ; Publication No. US2003005224A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gao, Wei-Oiang
 ; APPLICANT: Polakis, Paul
 ; APPLICANT: Shou, Jianyong
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Soriano, Robert
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wu, Thomas D.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
 ; TREATMENT OF TUMOR
 ; FILE REFERENCE: P5007R1-US

```

: CURRENT APPLICATION NUMBER: US/09/929,769
: CURRENT FILING DATE: 2001-08-14
: PRIOR APPLICATION NUMBER: 60/089,653
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/090,355
: PRIOR FILING DATE: 1998-06-23
: PRIOR APPLICATION NUMBER: 60/104,257
: PRIOR FILING DATE: 1998-10-14
: PRIOR APPLICATION NUMBER: 60/119,537
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: 60/141,037
: PRIOR FILING DATE: 1999-06-23
: PRIOR APPLICATION NUMBER: 60/145,698
: PRIOR FILING DATE: 1999-07-26
: PRIOR APPLICATION NUMBER: 60/162,506
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: PCT/US99/20111
: PRIOR FILING DATE: 1999-09-01
: PRIOR APPLICATION NUMBER: PCT/US99/28634
: PRIOR FILING DATE: 1999-12-01
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US00/00219
: PRIOR FILING DATE: 2000-01-05
: PRIOR APPLICATION NUMBER: PCT/US00/00376
: PRIOR FILING DATE: 2000-01-06
: PRIOR APPLICATION NUMBER: PCT/US00/04342
: PRIOR FILING DATE: 2000-02-18
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: PCT/US00/13705
: PRIOR FILING DATE: 2000-05-17
: PRIOR APPLICATION NUMBER: PCT/US00/22328
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: PCT/US01/06520
: PRIOR FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: PCT/US01/20118
: PRIOR FILING DATE: 2001-06-22
: PRIOR APPLICATION NUMBER: 09/888,257
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 8
: SEQ ID NO 3
: LENGTH: 3437
: TYPE: DNA
: ORGANISM: Homo Sapien
US-09-929-769-3

Alignment Scores:
Pred. No.: 0 Length: 3437
Score: 5241.00 Matches: 1029
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-929-769-7 (1-1029) x US-09-929-769-3 (1-3437)
QY 1 MethisileuValIhIsAlaMetValIleuLeuThrLeuGluProPArgAla 20
Db 46 ATGACATCTCTCGTGGCCATGCGATGCTGCTGACGCTGGCCCGCTCGAGCC 105
QY 21 AspAspSerGluPheGlnAlaLeuLeuAspIleTrpPheProGluGluysProLeuPro 40
Db 106 GAGCAGACGAGGTCCAGCGGCTGCTGACATCTGTTCCGAGGAGGAGCCACTGCC 165
QY 41 ThrAlaPheLeuValAspTrpSerGluGluAlaLeuLeuProAspTrpLeuysLeu 60
Db 166 ACCGCTTCTCTGTGGACATCGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 225
QY 61 ArgMetIleArgSerGluValIleuArgLeuValAspAlaIleuGlnAspLeuGluPro 80
|||||
```

```

Db 226 CGCATGATCCGTTCTGAGTGTCTCCGCTGTGGAGCCGCCCTGCAGACCTGAGCCG 285
QY 81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerIysLeu 100
Db 286 CAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
QY 101 LeuGlnPheLeuAspGlnAlaValAlaIhIsAspProGlnTrpLeuGlnGlnAsnIleMet 120
Db 346 CTCGAGTCTCTGAGCAGCAGGAGTGAGCCAGACCCCGAGCTGAGCAGACATCATG 405
QY 121 AspIysAsnTrpMetAlaIhIsLeuValGlnValGlnIhIsGluValArgIysAlaSerGly 140
Db 406 GACAGAGATTACATGAGCCCTGCTGAGGTGAGTCCAGCATGAGCCGCGCTCCGAGGC 465
QY 141 GlnThrPheIhIsSerLeuLeuThrAlaSerLeuProProArgAspSerThrGluAla 160
Db 466 CAGACTTTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
QY 161 ProIysProIysSerSerProGluGlnProIleGlyGlnIArgValGlyThr 180
Db 526 CCCAAGCCAAAGAGCAGCCAGCCAGCCAGCCATGAGCCAGGCGCGGATCCGGTGGGAGCC 585
QY 181 GlnLeuArgValLeuGlyProGluAspAspLeuValGlyMetPheLeuGlnIlePhePro 200
Db 586 CAGCTCCGAGTGTGAGCCCTGAGGAGCAGCTGCTGAGCATGTTCTCCATTTTCCG 645
QY 201 LeuSerProAspProArgTrpGlnSerSerSerProArgProValAlaLeuAlaLeuGln 220
Db 646 CTCAGCCCGGAGCCCTGAGTGGCAGAGCTCCAGCCCGGCGGCGGCTGCTGCTGCTGAG 705
QY 221 GlnAlaLeuGlyGlnGluLeuAlaArgValValGlnIysSerProGluValProGlyIle 240
Db 706 CAGCCCTGTGGCCAGAGCTGCGCCGCTGCTCCAGGAGCCAGGCTGCCGAGCTC 765
QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProIhIsGlyValAlaLeu 260
Db 766 ACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 261 ValMetSerMetIhIsArgSerIhIsPheLeuAlaCysProLeuLeuArgGlnLeuCysGln 280
Db 826 GTGATGTCCATGACCGGAGCCGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeuLeu 300
Db 886 TACAGCGCTGTGTGCCACAGACACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
QY 301 GlnMetLeuGlnTrpLeuAspSerProGlyValGlnGlyGlyProLeuArgAlaGlnLeu 320
Db 946 CAGATGCTGACGTGCTGAGACAGCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005
QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyIleu 340
Db 1006 AGGATGCTTGTGACCGCCCTGAGCGGCGGCGGCTGACTGATGTGCGAGGAGGAGGCT 1065
QY 341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGluValIleSerSerThrVal 360
Db 1066 CTGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1125
QY 361 ArgAlaValIleAlaTrpLeuArgSerGlyGlnIysSerValGluProAspLeuIle 380
Db 1126 CGTCCGCTGATCGGACACCTGAGGTCTGGGAGAGAGTGCAGACGAGGAGCGGACTGATC 1185
QY 381 SerIysValLeuGlnGlyLeuIleGluValArgSerProIhIsLeuGlnGluLeuThr 400
Db 1186 AGCAAGTCTCTCCAGGAGGCTGATGAGGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245
QY 401 AlaPhePheSerAlaTrpAlaAspAlaIleSerProPheProAlaCysIysProValVal 420
Db 1246 GCATTTCTTCTGTCCATGCGGAGTGTGCTCCCGCTTCCAGGCTGTAAACCGGTTGTG 1305
QY 421 ValValSerSerLeuLeuLeuGlnGluGluProLeuAlaGlyIysProGlyVala 440
Db 1306 GTGCTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1365
```

QY 441 AspGlySerLeuAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrp 460
 Db 1366 GAGGTGGACACCTGGAGCCGCTGGGCGCTGGTCAAGCTCTCTAGTGCAGC 1425
 QY 461 LeuGluMetLeuAspProGluValAlaSerSerCysProAspLeuGluLeuArgLeu 480
 Db 1426 CTGGAAATGCTGGACCCCGAGGTGGTCAAGCTGCCCCGACCTCCAGCTCAGCTGCTTC 1485
 QY 481 PheSerArgArgGlyGlyGlyGluAlaGlnValProSerPheArgProTrpLeuLeu 500
 Db 1486 TTCCTCCGAGAAAGGAGGAGGTCAGGCCAGGTGCCCTGCTGCTCCCTACCTCTG 1545
 QY 501 ThrLeuPheTrpHisGlnSerSerTrpProTrpHisGlnCysIleArgValLeuLeu 520
 Db 1546 ACCCTCTCAGCAGCATCTCCAGCTGGGCCACACAGCAGCATCGAGTCCGAGTCTGCTG 1605
 QY 521 GlyLysSerArgGluGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys 540
 Db 1606 GGCAGAGACCCGGGAAACAGAGGTTCCAGCCCTCTGCTGAGCTCTCTGCGGCTG 1665
 QY 541 IleHisValProArgIleTrpGlnArgAspGlnArgTrpProGlnLysArgArgGlu 560
 Db 1666 ATCCATGTTCTCCGATCTGGAGGGGCGGAGCACCGCACCCGAGAAAGCGCGGAG 1725
 QY 561 GluLeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGluLeuLeuAla 580
 Db 1726 GAGCTGGGTGGGCTGGGCTCAGGGCCCGGAGCTCATCAGCTGCTGAGTGTCTGCGC 1785
 QY 581 GluAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg 600
 Db 1786 GAGCGGAGAGCGCGGAGGACAGCGGAGACACAGCCGCTGACGCTCATCCAGGCGCG 1845
 QY 601 LeuProIleLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValThrGlu 620
 Db 1846 CTGCCCCCTGCTCTAGCTGCTGCTGGGAGCATGAGTGAAGAGGTGACGGAG 1905
 QY 621 HisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgArgCysArgAsp 640
 Db 1906 CACCTGTGAGGCTGCATCAGAGAGTGGGAGACACGCTGCTGGAAAGGCGTCCGAGAGC 1965
 QY 641 LeuLeuLeuGlnLeuTrpLeuGlnArgProGluLeuArgValProValProGluValLeu 660
 Db 1966 CTTCCTCTGAGCTCTACCTACAGCGGCGGAGCTGCGGCTGCCGTGAGGCTCTTA 2025
 QY 661 LeuHisSerGlnGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuLeuHisArg 680
 Db 2026 CTGCACAGCGAAGGGGCTGCCAGCAGCGCTCTGCAAGCTGGAGCTCATCCACCGC 2085
 QY 681 PheIleTrpLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnAsnArgGlyValAsp 700
 Db 2086 TTCATCAGCGCTCTTGGGAGACACAGCGACTCCGGGCTTGGAAACCGAGGGCGGAT 2145
 QY 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuArgHisLeu 720
 Db 2146 GCCACATAGGCTTCCGGAAGCTGGGAGCGGCGACCCGCTCTGCTCAGGACCTG 2205
 QY 721 PrometIleAlaAlaLeuLeuHisGlyArgTrpHisLeuAsnPheGlnGluPheArgGln 740
 Db 2206 CCCATGATCGGCGCTCTGACAGCGCGCACCCATCACTTCAGAGAGTTCCGGCG 2265
 QY 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGluLeuLeuInProHis 760
 Db 2266 CAGAACACACCTAGCTGCTCTGACAGCTGGGCGCTGCTGAGAGCTGTGACGCGCAC 2325
 QY 761 ValPheArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeu 780
 Db 2326 GTGTTCCCGACGAGCACACAGGGGCGGTGGAGAGCTCTGCTCTCATCCGCGCTG 2385
 QY 781 LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln 800
 Db 2386 CTGCTGAATTTACAGGAAGCTCTCCGCGCATCTGCTGCTCATCAACAAAGTTTGTGAG 2445

QY 801 PheIleHisLysTrpIleTrpTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHis 820
 Db 2446 TTCATCCATTAAGTACATTACCTACAAATGCCCCAGAGGCATCTCTCTGAGAAAGCAC 2505
 QY 821 AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu 840
 Db 2506 GCCGACCCGCTCCACAGACTCTCTTCGACAAACACTGACCTGTGATCTGTAATCCCTC 2565
 QY 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgTrpAspArgGlyLeuAspGlu 860
 Db 2566 CTTCGAGGCTCAGCTCTGCCACAGCAGGAGACAGACAGACCGAGCGCTGAGCAAGAG 2625
 QY 861 GlyGluGluGluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheTrpPro 880
 Db 2626 GCGGAGAGAGAGAGCTCAGCGCGCTCTGCCCCCTGGTACAGGTCTCTTACCCCT 2685
 QY 881 LeuThrAlaAlaGluMetAlaProTrpMetLysArgLeuSerArgGlyGlnThrValGlu 900
 Db 2686 CTGACCGCGCGCGAGATGAGCCCTTCATGAAACGCTTTCCCGGGGCCAAACGCTGAG 2745
 QY 901 AspleuLeuGluValLeuSerAspIleAspGluMetSerArgArgArgProGluLeuLeu 920
 Db 2746 GATCTGCTGAGAGTCTGTAGTACATACAGCATGATCCCGGCGAGACCGAGATCTG 2805
 QY 921 SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGluGluCysArgAsn 940
 Db 2806 ACCTTCTCTCTGACCACTGACGCGCTGATGAGCTGCGGCGAGAGTGTGCCGCAAC 2865
 QY 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaPhe 960
 Db 2866 CTGCGCTTCAGCTGCGCTGCGCTTCATGCAAGAACGCGCCAGCATTTGACGCCCTTTC 2925
 QY 961 LeuProThrPheMetTrpCysLeuGlySerGlnAspPheGluValAlaGlnThrAlaLeu 980
 Db 2926 CTGCCACGTTATATGATCTGCTGGGACAGCAGGACTTTGAGAGTGTGACAGCGGCTC 2985
 QY 981 ArgAsnLeuProGluTrpAlaLeuLeuCysGlnGluHisAlaAlaValLeuLeuHisArg 1000
 Db 2986 CGGAACCTGCGCTGAGTACGCTCTCTGTCGCAAGAGACGCGGCTGTGCTCCACCGG 3045
 QY 1001 AlaPheLeuValGlyMetTrpGlyGlnMetAspProSerAlaGlnIleSerValAlaLeu 1020
 Db 3046 GCCTTCTGCTGGGATGATGACGCGCAGATGAGCCACCGCGCAGATCTCCAGGCGCTG 3105
 QY 1021 ArgIleLeuHisMetGlnAlaValMet 1029
 Db 3106 AGGATCTGCAATATGAGAGCGGCTGATG 3132

RESULT 8
 US-10-032-996-21
 ; Sequence 21, Application US/10032996
 ; Publication No. US20030054447A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Geo, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Pan, James
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tamas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2930R1C3
 ; CURRENT FILING DATE: 2001-12-27
 ; PRIOR APPLICATION NUMBER: 60/095,325
 ; PRIOR FILING DATE: 1998-08-04


```

|||||
Db 1306 GTGGAGACCTCCCTGCTCTGAGAGAGAGAGCCCTGCGTGGGGAGAGCCGCTGCG 1365
Qy 441 ASpglySerLeuGluAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrp 460
Db 1366 GACGGTGGACACCTGGAGAGCGGTGGGCGCTGCTGACGGCTCTCTAGTGAGACGG 1425
Qy 461 LeuGluMetLeuAspProGluValValSerSerCysProAspLeuGluLeuArgLeuLeu 480
Db 1426 CTGGAATAGCTGGAGCCCGAGAGTGTGACAGACTGCCCGAGACTGACAGCTGCTGCTC 1485
Qy 481 PheSerArgArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 500
Db 1486 TTCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1545
Qy 501 ThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeuLeu 520
Db 1546 ACCCTCTTCACGACATCTCCAGCTGGGCCACACTGCACAGTCGATCCGACTCTGCTG 1605
Qy 521 GlySerSerArgGluGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys 540
Db 1606 GGCAAGACCCGGGAACAGAGGTGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1665
Qy 541 IleHisValProArgIleTrpGlnGlyArgAspGlnArgTrpProGlnLysArgArgGlu 560
Db 1666 ATCATGTTCTCCGATCTGGCAGAGGGCGGAGCACAGCCACCCCGACAGAACGGCGGAG 1725
Qy 561 GluLeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGluLeuIleLeuAla 580
Db 1726 GAGCTGGTGTGGGGGCTGACAGGGCCGAGCTCATCAGCTGCTGCTGCTGCTGCTGCTG 1785
Qy 581 GluAlaGlnTrpArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg 600
Db 1786 GAGGGGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1845
Qy 601 LeuProLeuLeuLeuSerCysCysGlyAspAspGluSerValArgLysValThrGlu 620
Db 1846 CTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1905
Qy 621 HisLeuSerGlyCysIleGlnGlnTrpGlnAspSerValIleuGlyArgArgCysArgAsp 640
Db 1906 CACCTGTAGAGCTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1965
Qy 641 LeuLeuLeuGlnLeuTrpLeuGlnArgProGluLeuArgValProValProGluValLeu 660
Db 1966 CTCTCTCTGACGCTCTACTACAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 2025
Qy 661 LeuHisSerGluGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuIleHisArg 680
Db 2026 CTGCACACCGAAGGGGCTGCCAGCAGCAGCGCTGCCAAGCTGGAGAGACTCATCCACCGC 2085
Qy 681 PheIleThrLeuLeuAlaAspTrpSerAspSerArgAlaIleuGlnAspArgGlyAlaAsp 700
Db 2086 TTCTATCAGAGCTCTTGGCGAGACACAGCACTCCCGGGCTTGGAGAACCGAGGGCGGAT 2145
Qy 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuArgHisLeu 720
Db 2146 GCCAGCAGAGCTGCTGCCGAGAGTGGCGGTGGCGGACCCGCTGCTGCTGCTGCTGCTGCTG 2205
Qy 721 PrometIleAlaAlaLeuLeuHisGlyArgThrHisLeuAspPheGlnGluPheArgGln 740
Db 2206 CCCATGATCGGGGCTCTCTCAGCGCGCACCCCACTTCACAGAGAGTTCGGGAGAG 2265
Qy 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGlnProHis 760
Db 2266 CAGAACCCAGCTGAGCTCTTCTGACAGCTGCGGGCTGCTGCTGCTGCTGCTGCTGCTG 2325
Qy 761 ValPheArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeu 780
Db 2326 GTGTGTCCGACGAGACACAGAGGGGCGCTGTGGAGAGCTGCTGCTGCTGCTGCTGCTG 2385
Qy 781 LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaIlePheIleAsnLysPheValGln 800
|||||

```

```

Db 2386 CTGCTGAATTACAGGAAGTCTCCCGCATCTGAGCTGCTTCATCACAGATTGTCGAG 2445
Qy 801 PheIleHisLysTrpIleThrTyraAsnAlaProAlaAlaIleSerPheLeuGlnLysHis 820
Db 2446 TTCTATCTCAATAGTACATTACCTACAAATGCCCGAGAGCCATCTCTTCTGACAGAGAC 2505
Qy 821 AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu 840
Db 2506 GCCGACCCGCTCCAGAGACTGCTCTTCGACAAACATGACCTGCTGATGCTGAATCCCTTC 2565
Qy 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgThrAspArgGlyLeuAspGluGlu 860
Db 2566 CTTCGAGGGCTTCAGCTGCTCCAGCAGGAGCAGCAGAGACCGACCGAGCTGGACGCAAGAG 2625
Qy 861 GlyGluGluGluSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrPro 880
Db 2626 GCGGAGAGAGAGAGCTCAGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2685
Qy 881 LeuThrAlaAlaGluMetAlaProTrpMetLysArgLeuSerArgGlyGlnThrValGlu 900
Db 2686 CTGACCGCGCGCGAGATGGCCCTACATGAAAGGCTTTCGCGGGCCAAACGGTGGAG 2745
Qy 901 AspLeuLeuGluValLeuSerAspIleAspGluMetSerArgArgArgProGluIleLeu 920
Db 2746 GATCTGTGTGAGGTTCTGATGATACATAGACGAGATGTCGCGGAGACCCGAGATCCTG 2805
Qy 921 SerPhePheSerThrAsnLeuGlnArgProLeuMetSerSerAlaGluGluCysArgAsn 940
Db 2806 AGCTTCTTCTGACCAACCTGACGCGCTGATGAGACTGCGGCGAGAGAGTGTGGCAAC 2865
Qy 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaIlePhe 960
Db 2866 CTGCGCTTCAGGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2925
Qy 961 LeuProThrPheMetCysLysLeuGlySerGlnAspPheGluValGlnIleAlaLeu 980
Db 2926 CTGCCCCAGTTTCAATGATCTGCTGCGGAGCCAGGACTTGTGAGTGTGTGACAGCGCCCTC 2985
Qy 981 ArgAsnLeuProGluTrpAlaLeuLeuCysGlnGlnIleAlaAlaValLeuLeuHisArg 1000
Db 2986 CGGAACCTGCTGAGTACGCTTCTGCTGCCAAGAGCAGCGGCTGTGCTGCTCCACCGG 3045
Qy 1001 AlaPheLeuValGlyMetTrpGlyGlnMetAspProSerAlaGlnIleSerGluAlaLeu 1020
Db 3046 GCCTTCTGTGTGGGATGATGAGGCGCAGATGAGACCCAGGCGCAGATCTCCGAGGCCCTG 3105
Qy 1021 ArgIleLeuHisMetGluAlaValMet 1029
Db 3106 AGGATCTGCAATATGAGGCGGTGATG 3132

RESULT 9
US-10-033-396-21
; Sequence 21, Application US/10033396
; Publication No. US2003007657A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2930R1C4
; CURRENT APPLICATION NUMBER: US/10/033, 396
; CURRENT FILING DATE: 2001-12-27

```

Dp	166	ACCSCCTCCGTGGGACACATCGAGGAGAGGCGCTCTCTTCCTGACGAGCTGAAGCTG	225
Qy	61	ArgMetIleArgSerGluValIleuAlaGleuValAspAlaIaIaIeGlnAspLeuGluPro	80
Dp	226	CGCATGTATCCGTTCTGAGGTGCTCCCTCGTGTGGACGCGCCCTTCAGAGACTTGAGGCTG	285
Qy	81	GlnIleuLeuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerIysLeu	100
Dp	286	CAGAGCTGCTGCTGTTCGGAGCTGTTGGCATGCCGTGTCCAGATGAGCAACTTC	345
Qy	101	LeuGlnPheLeuAspGlnAlaValAlaIaHisAspProGlnThrLeuGlnGlnAsnIleMet	120
Dp	346	CTCCAGTTCCTGGACCCAGGACATGGGCCACAGACCCCAACACTCTGGAGCAGAACATCTAG	405
Qy	121	AspIysAsnIlyrMetAlaHisLeuValGluValGlnHisGluArgGluValAserGlyIle	140
Dp	406	GACAGAAATTTACATGAGGCCCACTGCTGTGGAGTCCAGACATGAGCGGGCCCTCCGAGAC	465
Qy	141	GlnThrPheHisSerLeuLeuThrAlaSerLeuProProArgAlaAspSerThrGluAla	160
Dp	466	CAGACTTTCACATCTTGTCTCACAGCCTCCCTGCGCCCGCGAGACAGCACAGACGCA	525
Qy	161	ProIysProIysSerSerProGluGlnProIleGlyGlnGlyArgIleLeuValGluThr	180
Dp	526	CCCAACACCAAGAGACACCCAGACAGCCCATAGCCACAGGCGCGATTCGGGTGGGAGCC	585
Qy	181	GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro	200
Dp	586	CAGCTCGGGGCTGGGACCTCGAGAGACAGACTGGCTGGCATTTCTCCAGATTTTCCCG	645
Qy	201	LeuSerProAspProAlaGlyTrpGlnSerSerSerProArgProValAlaLeuAlaLeuGln	220
Dp	646	CTTCAGCCCGGACCTCGGTGGACAGCTCCAGTCCCGCCCGCTGGCCCTGCTGCTGACG	705
Qy	221	GlnAlaLeuGlyGlnGluLeuAlaArgValValGlnGlySerProGluValProGlyIle	240
Dp	706	CAGGCCCTGGGGCCAGAGAGCTGGCCGGCTGTCAGAGGACGCCGAGAGTGGCCGGGATC	765
Qy	241	ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyIleValLeu	260
Dp	766	ACGGTGGGTCTGTCGACAGGCCCTCGCACCTGCTACTGCCCAACAGCGCGTCTCCCTG	825
Qy	261	ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgIleuLeuGlyGln	280
Dp	826	GTGATGTCATGACACCGTACCACACTTCCTGCGTCCGCTGCTGGCCAGCTCTGCCAG	885
Qy	281	TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuIysValLeuLeu	300
Dp	886	TACCAGCGCTGTGGCCACAGACACCGGCTTCTCTGCTCTTCTGTAAGGTGCTCTCTG	945
Qy	301	GlnMetLeuGlnTrpLeuAspSerProGlyValGlnGlyIleProLeuArgIleGlnLeu	320
Dp	946	CAGATGCTGCAGTGGCGAGACACCCCTGGGCTGTGAGGCGGGCCCTCGGGGACACAGCTC	1005
Qy	321	ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyIleLeu	340
Dp	1006	AGGATGCTTGCACACCGAGGCTCAGCGCGGGCGAGGCTCAGTGAATGTGGAGGGGGCTC	1065
Qy	341	LeuArgLeuAlaGluAlaLeuAlaPheArgGlnAspLeuGluValIaIaSerSerThrVal	360
Dp	1066	CTGCGGCTGGCCGAGGCGCTCGGCCCTTCCTGACAGACTGAGAGTGGTGAAGCTCCACGCTC	1125
Qy	361	ArgAlaValIleAlaThrLeuArgSerGlyGlnGlnCysSerValGluProAspLeuIle	380
Dp	1126	CTGCTCCGTCATCGCACCTGAGGCTGGGGAGACATGACAGGTGGAGCGGACCTGTATC	1185
Qy	381	SerIysValLeuGlnGlyLeuIleGluValArgSerProHisLeuGlnGluIleuLeuThr	400
Dp	1186	AGCAAAAGTCCTCCAGGGGCTGATCGAGGTGAGGTGCCCCACACTGAGAGAGCTGCTGACT	1245
Qy	401	AlaPhePheSerAlaThrAlaAspAlaIaIaSerProPheProAlaCysIysProValVal	420
Dp	1246	GCATTTCTTCTGTGCATGTGGATGTGCTCCCGCTTTCACGCTGTAAGCCCGTTGTG	1305

[illegible]

QY	781	LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln	800
Db	2386	CTGGCAATTACAGAAAGTCTCCCGCATCTGGCTGCTTATCAACAAGATTCTGCAG	2445
QY	801	PheIleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHis	820
Db	2446	TTTCATCCATTAGTACATTACCCTCAATAGCCCCAGCAGCATCTCCCTTCCTGCAGAAAGCAC	2505
QY	821	AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu	840
Db	2506	GCCGACCCCGCTCCACGACGCTGCTCTTCGCAACAAGAGCACTGGTATGCTGAAATCCCTC	2565
QY	841	LeuAlaGlyLeuSerLeuProSerArgAspAspArgTrpAspArgGlyLeuAspGluGln	860
Db	2566	CTTCGACGGGCTCAACCTCGCCACAGAGAGACAGACAGCCAGCAGGCTTGACCAACAAG	2625
QY	861	GlyGluGluGlnSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrPro	880
Db	2626	GCGCAGGAGGAGACCTCAGCCGGCTCTCTTCCCTGGTGAAGCTCTCCCTGTTCACCCCT	2685
QY	881	LeuThrAlaAlaGluMetAlaProTrpMetLysArgLeuSerArgGlyGlnThrValGlu	900
Db	2686	CTGACCCGGGCGGAGATGGCCCCCTACATGAAACGGCTTTCCCGGGCCCAACGGTGGAG	2745
QY	901	AspLeuLeuGluValLeuSerAspIleAspGluMetSerArgArgArgProGluIleLeu	920
Db	2746	GATCTGCTGAGGAGTTCTGAGTGCATGACATGAGACAGATGTCCCGCGAGACCCAGCATCTG	2805
QY	921	SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGluGlnLysArgAsn	940
Db	2806	AGCTTCTTCTCGACCAACGTCGACGGCTGATAGTCTGGCGAGAGAGTGGCCGAC	2865
QY	941	LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaAlaPhe	960
Db	2866	CTCGCCTTCACACCTGGGCGCTCCATGCATCAACAACGCCACGATTTGACCCGCTTTC	2925
QY	961	LeuProThrPheMetTrpCysLeuGlySerGlnAspPheGluValValGlnThrAlaLeu	980
Db	2926	CTGCCACAGTCATGACTCTCCCTGGGAGCGACAGACTTTGAGAGTGGTCAGACGGCCCTC	2985
QY	981	ArgAsnLeuProGluTrpAlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArg	1000
Db	2986	CGGAACCTGCTGAGTACGCTCTCTGTGCCAAGACAGCGCGGTGTGCTGCCACGG	3045
QY	1001	AlaPheLeuValGlyMetTrpGlyGlnMetAspProSerAlaGlnIleSerGluAlaLeu	1020
Db	3046	GCTCTCTCGGTGGGCATGTACGGCGCAGATGGACCCAGCGCGCATCTCCAGAGCCCTG	3105
QY	1021	ArgIleLeuHisMetGluAlaValMet	1029
Db	3106	AGGATCTCGCATATGAGAGGCCGTGATG	3132
RESULT 10			
US-10-033-246-21			
: Sequence 21, Application US/10033246			
: Patent No. US20020098505A1			
: GENERAL INFORMATION:			
: APPLICANT: Botsstein, David			
: APPLICANT: Desnoyers, Luc			
: APPLICANT: Ferrara, Napoleone			
: APPLICANT: Fong, Sherman			
: APPLICANT: Gao, Wei-Qiang			
: APPLICANT: Goddard, Audrey			
: APPLICANT: Gunney, Austin L.			
: APPLICANT: Pan, James			
: APPLICANT: Roy, Margaret Ann			
: APPLICANT: Stewart, Timothy A.			
: APPLICANT: Tumas, Daniel			
: APPLICANT: Watanabe, Colin K.			
: APPLICANT: Wood, William I.			
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic			
: TITLE OF INVENTION: Acids Encoding the Same			
: FILE REFERENCE: P2930R1C12			

```

CURRENT APPLICATION NUMBER: US/10/033,246
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 21
LENGTH: 3437
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-246-21

Alignment Scores:
Pred. No.: 0 Length: 3437
Score: 5241.00 Matches: 1029
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-929-769-7 (1-1029) x US-10-033-246-21 (1-3437)
Oy 1 MethistileuvalValhisAlametaVallleuleuthreuglyProPofarGala 20
|||||
Db 46 ATGCACATCCCTCGGTGCCATCATGGTATCCTCGTGAGCGGTGGCCCCGCTCGAGCC 105
|||||
Oy 21 AspaspsergiUphegInAlaleuleuaspllerppheprognigugluyLysProleupro 40
|||||
Db 106 GACGACGCGAGTTCCAGGCGCCCTCTGACATCTGTGGTTTTCCGGAGGAGAAAGCCACTGCCC 165
|||||
```

QY	41	ThlAlpheleuValAspThrSerGluGluAlaLeuLeuLeuProAspTyrLeu	60
Db	166	ACCGCTTCCTGGTGGACATCGGAGAGGCGCTGCTTCCAGTGGTAAGCTG	225
QY	61	ArgMetIleArgSerGluValLeuArgLeuValAspAlaIleLeuGlnAspLeuGlnPro	80
Db	226	CGCATGATCCGTTCTGAGAGTCTCCGCTGGTGAGCCGCCCTCGAGAGACCTGAGCCG	285
QY	81	GlnGlnIleuLeuLeuPheValGlnSerPheGlyTyrProValSerSerMetSerLysIleu	100
Db	286	CAGACGCTGCGTGGTTCGTGCACTCTTTGGCATCCCGCTGTCAGCAAGACAAATC	345
QY	101	LeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGlnGlnAsnIleMet	120
Db	346	CTCCAGTTCCTGGACACAGGAGAGTGGCCACGAGCCCAAGCTCTGGAGCAGAACATCATG	405
QY	121	AspLysAsnTyrMetAlaHisLeuValGluValGlnHisGluArgIleValSerGlyGly	140
Db	406	GACAAGATTATCATGGCCCACTGTTGAGATGCAGATGAGCGGCGCTCCGGAGGC	465
QY	141	GlnThrPheHisSerIleuLeuThrAlaSerLeuProProArgArgAspSerThrGlnAla	160
Db	466	CAGACTTTCCTCCTCTGCTACAGCCCTCCCTGCGCCGCGGAGACACACAGAGGCA	525
QY	161	ProLysProLysSerSerProGlnGlnProIleGlyGlnArgIleArgValGlyThr	180
Db	526	CCCAACACCAAGACAGACGCCAGACAGCCCATAGGCCAGGCGCGGATTCGGTGGGAGCC	585
QY	181	GlnIleuArgValIleuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro	200
Db	586	CAGCTCCGGGTGGGGGCCCTCGAGAGAGCACTGGCTGGCATGTTCTCCAGATTTTCCG	645
QY	201	LeuSerProAspProArgTyrGlnSerSerSerProArgProValAlaLeuAlaLeuGln	220
Db	646	CTCAGCCCGAGACCTCGGTGGCAAGCTCCAGTCCCGGCCGCGGGCCCTCGCGAG	705
QY	221	GlnAlaLeuGlyGlnGlnIleuAlaArgValValGlnGlySerProGluValProGlyIle	240
Db	706	CAGCGCCCTGGGCGAGAGAGCTGGCCGGTGTCTCCAGGCGAGCCCGAGGTGGCGGATC	765
QY	241	ThrValArgValIleuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyGlyAlaLeu	260
Db	766	ACGGGCTGTCTGCGAGAGCCCTGCCACCTGCTACGCTCCCAACAGGCGGTGCCCTG	825
QY	261	ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGln	280
Db	826	GTGATGTCATGACACCGAGCACTTCTGGCTGGCCGCTGCTGGCGCCAGCTCTCCAG	885
QY	281	TyrGlnArgCysValProGlnAspThrGlyPheSerSerIleuPheLeuLysValIleuLeu	300
Db	886	TACCAAGCCTGTGTGCCACAGACACCGGCTTCTCTCTGCTTCTCGAAGGTGCTCTG	945
QY	301	GlnMetLeuGlnTyrPheuAspSerProGlyValGlnGlyGlyProLeuArgAlaGlnLeu	320
Db	946	CAGATGCTGAGTGGTGGAGACGCCCTGGGCTGGAGGGCGGCGCCCTGCGCGACAGCTC	1000
QY	321	ArgMetIleuAlaSerGlnAlaSerAlaGlyValArgArgIleuSerAspValArgGlyGlyIleu	340
Db	1006	AGGATGCTTGGCACCGAGCTCAGCGGGGCGAGGCTCTAGTATGTGCGAGGGGGGCTC	1060
QY	341	LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGluValValSerSerThrVal	360
Db	1066	CTGGCGCTGGCCGAGGCGCTGGCTTCGTCAGAGCACTGGAGGGTCACTCCACGCTC	1120
QY	361	ArgAlaValIleAlaThrLeuArgSerGlyGlnGlnCysSerValGlnProAspLeuIle	380
Db	1126	CGTGGCGTCATCGCCACCTGATGAGTCTGGGAGACAGTGCAGCTGGAGCGCGACCTGATC	1180
QY	381	SerLysValIleuGlnGlyLeuIleGlyValArgSerProHisLeuGlnGlnIleuLeuThr	400
Db	1186	AGCAAAAGCTCTCCAGGGGCTATGAGAGTGAAGTCCCCCACTGGAGAGAGCTGCTGACT	1240
QY	401	AlaPhePheSerAlaThrAlaAspAlaAlaSerProPheProAlaCysLysProValVal	420

```

|||||
Db 1246 GCATCTCTCTGCGACATGCGGATGCTGCTCCCGTTCCAGCTGTAAGCCGCTGG 1305
Qy 421 ValValSerSerLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 440
Db 1306 GTGGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1365
Qy 441 AspGlySerLeuGlnAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrp 460
Db 1366 GACGGTGACCTGTGAGGCGCTGGGCGCTGGGCGCTGTCAAGCGCTCTAGTGAAC 1425
Qy 461 LeuGlnMetLeuAspProGlnValValSerSerCysProAspLeuGlnLeuArgLeu 480
Db 1426 CTGGAAAGCTGACCCCGAGGTGGTCAAGAGCTCCCGACCTCAGCTCAGCTGCTC 1485
Qy 481 PheSerArgArgGlyGlyGlyGlnAlaGlnValProSerPheArgProTrpLeuLeu 500
Db 1486 TTCTCCCGAGAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1545
Qy 501 ThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeu 520
Db 1546 ACCCTCTCAGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1605
Qy 521 GlySerSerArgGlnGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys 540
Db 1606 GGCAGAGACCGGGAACAGAGGTTGACCCCTCTGCTCTGAGACTCTCTGCGGCTG 1665
Qy 541 IleHisValProArgIleTrpGlnGlyArgAspGlnArgTrpProGlnLysArgArg 560
Db 1666 ATCCATGTTCTCGCATGTGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1725
Qy 561 GluLeuValLeuArgValGlnGlyProGlnLeuLeuSerLeuValGlnLeuLeuAla 580
Db 1726 GAGCTGGTGTGGGGGCTCAGAGGCGGAGCTCATCATCATCATCATCATCATCAT 1785
Qy 581 GluAlaGlnTrpArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg 600
Db 1786 GAGGGGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1845
Qy 601 LeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValThrGln 620
Db 1846 CTGCCCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 1905
Qy 621 HisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgArgCysArgAsp 640
Db 1906 CACCTGTGAGCTGCTCATCAGAGTGGGAGAGACAGCTGCTGGAGGGGCTGCCGAGAC 1965
Qy 641 LeuLeuLeuGlnLeuTrpLeuGlnArgProGlnLeuArgValProValProGlnValLeu 660
Db 1966 CTCTCTCTGAGCTCTACCTACAGGCGGAGCTGCGGCTGCTCTCTCTCTCTCTCT 2025
Qy 661 LeuHisSerGlnGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuLeuHisArg 680
Db 2026 CTGCACACGAGAGGGGCTGCCAGCAGACCTGTCTCAAGCTGTGAGGACTCATCAAC 2085
Qy 681 PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnAsnArgGlyAlaAsp 700
Db 2086 TTCAATCAGCTCTCTGCGAGACACAGCTCTCCGGGCTGGAGAACCCAGAGGGCGGAT 2145
Qy 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuLeuArgHisLeu 720
Db 2146 GCCACACAGGCTGCGGAGAGTGGCGTGGCGACCCGCTCTCTCTCTCTCTCTCTCT 2205
Qy 721 PrometIleAlaAlaLeuLeuHisGlyArgThrHisLeuAsnPheGlnGlnPheArgGln 740
Db 2206 CCCATGATCGGGGCTCTCTGACGCGGCGACCCACTCAACTTCAGAGAGTTCGGGCGAG 2265
Qy 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGlnProHis 760
Db 2266 CAGAACCACTGAGCTGCTCTCTGACAGTGGCGCTGCTGAGACTGCTGACGCGGAC 2325
Qy 761 ValPheArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeu 780
|||||

```

```

Db 2326 GTTTCGCGAGCGACACAGGAGGCGCTGTGAGACTGCTCTCTCTCTCTCTCT 2385
Qy 781 LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln 800
Db 2386 CTGCTGAATTAAGAGAGGCTCCCGCATCTGGGCTCTTATCAACAAGTTTGTGACG 2445
Qy 801 PheIleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHis 820
Db 2446 TTCATCCATTAAGTACATTAACCTACCAATGCCAGCAGCATCTCTCTCTGACAGAAC 2505
Qy 821 AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu 840
Db 2506 GCCGAGCCGCTCCAGACCTGCTCTGCAACAAGTGAACCTGTATGCTGTAATCCCTC 2565
Qy 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgTrpHisArgGlyLeuAspGlnGln 860
Db 2566 CTTCGAGGGCTGAGCTGCGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2625
Qy 861 GlyGlnGlnGlnSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrPro 880
Db 2626 GCGGAGAGAGAGAGCTCAGCGCGCTCTGCGCTGCTGAGGCTCTCCCTGTCAACCT 2685
Qy 881 LeuThrAlaAlaGlnMetAlaProTrpMetLysArgLeuSerArgGlyGlnThrValGln 900
Db 2686 CTGACCGCGCGCGAGATGCGCCCTACATGAACGCGTTTCCCGGCGCAACGCGTGAG 2745
Qy 901 AspLeuLeuGlnValLeuSerAspIleAspGlnMetSerArgArgArgProGlnLeuLeu 920
Db 2746 GATCTGCTGAGGTTCTGTGTACATAGCAGAGATGTCCCGGAGGAGCCGAGATCTCG 2805
Qy 921 SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGlnGlyCysCysArgAsn 940
Db 2806 ACCTCTCTCTCGACCACTGACGCGGTGATGAGCTGCGCGAGAGAGTGGCGGCAAC 2865
Qy 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaPhe 960
Db 2866 CTGGCTCTTCAAGCTGCGCTGCGCTCATGACAGAACAGCCCGACATTCACCCCTTTC 2925
Qy 961 LeuProThrPheMetTrpCysLeuGlnArgLeuGlnAspPheGlnValValGlnThrAlaLeu 980
Db 2926 CTGCCACGTTTATGATGCTGCTGGGACCGCAGACTTGTAGAGTGTGTGACAGCGCC 2985
Qy 981 ArgAsnLeuProGlnTrpAlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArg 1000
Db 2986 CGSAGACTGCTGAGTACGCTCTCTGCTCAAGAGACAGCGGCTGTGCTCCACCGG 3045
Qy 1001 AlaPheLeuValGlyMetTrpGlyGlnMetAspProSerAlaGlnIleSerValAlaLeu 1020
Db 3046 GCTTCTCTGGTGGGATGTACGGCCAGATGAGACCCAGCGCGCATCTCCAGGCGCTG 3105
Qy 1021 ArgIleLeuHisMetGlnAlaValMet 1029
Db 3106 AGGATCTCTCATATGAGGCGCTGATG 3132

```

RESULT 11

US-10-033-301-21

Sequence 21, Application US/10033301

Patent No. US20020098506A1

GENERAL INFORMATION:

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Goddard, Audrey

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

```

: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2930R1C6
: CURRENT APPLICATION NUMBER: US/10/033,301
: CURRENT FILING DATE: 2001-12-27
: PRIOR APPLICATION NUMBER: 60/095,325
: PRIOR FILING DATE: 1998-08-04
: PRIOR APPLICATION NUMBER: 60/112,851
: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,145
: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,511
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: 60/115,558
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/115,565
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/115,733
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/119,341
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: 60/119,537
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: 60/119,965
: PRIOR FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: 60/162,506
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: 60/170,262
: PRIOR FILING DATE: 1999-12-09
: PRIOR APPLICATION NUMBER: 60/187,202
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: PCT/US99/28634
: PRIOR FILING DATE: 1999-12-01
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US00/03565
: PRIOR FILING DATE: 2000-02-11
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: 2000-03-02
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: PCT/US00/14941
: PRIOR FILING DATE: 2000-05-30
: PRIOR APPLICATION NUMBER: PCT/US00/15264
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: 2000-12-01
: NUMBER OF SEQ ID NOS: 38
: SEQ ID NO 21
: LENGTH: 3437
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-033-301-21

Alignment Scores:
Pred. No.: 0 Length: 3437
Score: 5241.00 Matches: 1029
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-929-769-7 (1-1029) x US-10-033-301-21 (1-3437)
QY 1 MeHhSiIleuValYalHhSaIaMeTValIleuLeuThreugLyPProArGaLa 20
Db 46 ATGCACATCTCTGCTGCATGCCATGCTATCTCTGCGCGCCGCGCTCGAGCC 105
QY 21 AspAsSerGluPheGlnAlaLeuAspIleTirPheProGluGluYProLeuPro 40
|||||

```

```

Db 106 GACGACGCGAGTTCACGCGCTGCTGACATCTGCTTCGAGAGAGACCATGCCCC 165
QY 41 ThAlaPheLeuValAsPThrSerGluGluAlaLeuLeuProAsPThrLeuYLeu 60
Db 166 ACCGCTTCTGCTGGACATCGAGAGAGCGCTGCTGCTCTGCTGCTGAGCTG 225
QY 61 ArgMetIleArgSerGluValLeuArgLeuValAspAlaAlaLeuGlnAspLeuPro 80
Db 226 CGCATGATCCGTTCTGAGGTCCTCCGCTGTCGAGCGCCCTGACAGACCTGAGCGC 285
QY 81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProValSerMetSerLeu 100
Db 286 CAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
QY 101 LeuGlnPheLeuAspGlnAlaValAlaHhAspProGlnThreugLuglnAsnIleMet 120
Db 346 CTCACGTTCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 405
QY 121 AspLysAsnTyrMetAlaHhIleValGluValGlnHhIleGluArgGlyAlaSerGly 140
Db 406 GACAGAAATACATGCGCCACCTGCTGAGGTCACATGAGCGCGCGCTCCGAGGCG 465
QY 141 GlnThrPheHhSerLeuLeuThrAlaSerLeuProProArgAsPThrGluAla 160
Db 466 CAGACTTTCACCTCTTCTGCTCAGAGCTCCCTCCGCGCGCGAGACAGAGGCA 525
QY 161 ProLysProLysSerSerProGluGlnProIleGlyGlnIleArgValGlyThr 180
Db 526 CCCAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
QY 181 GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro 200
Db 586 CACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
QY 201 LeuSerProAsPProArgTirPginSerSerSerProArgProValAlaLeuAlaLeuGln 220
Db 646 CTCAGCGCGAGACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
QY 221 GlnAlaLeuGlyGlnLeuAlaArgValValGlnGlySerProGluValProGlyIle 240
Db 706 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHhGlyValAlaLeu 260
Db 766 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 261 ValMetSerMetHhIleArgSerHhIlePheLeuAlaCysProLeuLeuArgGlnLeuCysGln 280
Db 826 GTGATGTCATGACAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeuLeu 300
Db 886 TACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
QY 301 GlnMetLeuGlnThrLeuAspSerProGlnValGlnGlyProLeuArgAlaGlnLeu 320
Db 946 CAGATGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeu 340
Db 1006 AGGATGCTTCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1065
QY 341 LeuArgLeuAlaGluAlaLeuAlaPheArgGlnAspLeuGluValValSerSerThrVal 360
Db 1066 CTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
QY 361 ArgAlaValIleAlaThrLeuArgSerGlyGlnGlyCysSerValGluProAspLeuIle 380
Db 1126 CGTGCCTGATCCACACCTCGAGGTCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1185
QY 381 SerLysValLeuGlnGlyLeuIleGluValArgSerProHhIleGlnGluLeuLeuThr 400
Db 1186 AGCAAGTCTCTCAGGAGGCTGATGAGTGAAGTCCCGCCACCTGAGAGGCTGCTGACT 1245

```

OY	401	AlaPhePheSerLarThrAlaAspAlaAlaSerProPheProAlaCysLysProValVal	420
Db	1246	GCATTCTTCTGCGACGTGGAGTGTGCTCCCGATTCCAGCTGAACCCGTTGTG	1305
OY	421	ValValSerSerLeuLeuLeuGlnGlnGlnGlnProLeuAlaGlyLysProGlyAla	440
Db	1306	GTGGTGAAGTCCTCGTCTGCGACAGAGAGAGAGACCCTGCTGGGGGGAGCCGGGTGC	1365
OY	441	AspGlyLysSerLeuGlnAlaValAlaGlyLeuGlyProSerSerGlyLeuLeuValAspTrp	460
Db	1366	GAGCGTGGACGCTGGAGGGCGGTGGGCGTGGGGCCCTGTCAGAGGCTCTTATGTGACATGG	1425
OY	461	LeuGlnMetLeuAspProGlnValValSerSerCysProAspLeuGlnLeuArgLeuLeu	480
Db	1426	CTGGAATATGCTGACACCCCGAGGTGGTTCAGACGTGCGCCCGACACTCAGCTGCTGC	1485
OY	481	PheSerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTrpLeuLeu	500
Db	1486	TTTCTCCGGGAGAAAGGGCAAAAGGTCTCAGGCCCAAGTGTCCCTCTTCCTGCTTACCTCTG	1545
OY	501	ThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeuLeu	520
Db	1546	ACCTCTTTCACGCATCAGTCCAGCTGGCCACACTGACACATCGCATCCGAGTCTGCTG	1605
OY	521	GlyLysSerArgGlnGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys	540
Db	1606	GGCAAGAGCCGGGAAACAGAGAGTTCCAGACCCTTGCTCTGCTGACTTCCTGGGGCTGC	1665
OY	541	IleHisValProArgIleTrpGlnGlnArgAspGlnArgThrProGlnLysArgArgLys	560
Db	1666	ATTCATGTTCCTCGCACATCTGGCAGAGGGGGGAGCACAGCCGCCAGAGAAGGGGGAG	1725
OY	561	GluLeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGlnLeuLileuAla	580
Db	1726	GAGCTGGGTGCGCGGGTCCAGGGGCCCGAGACTATAGCGCTGTGGACTGATCTTGCC	1785
OY	581	GluAlaGluThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg	600
Db	1786	GAGCGCGAGACGCGAGCCAGGAGCGGGGACACAGCCGCTGCAGCCTCATCCAGGCCGG	1845
OY	601	LeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValThrLys	620
Db	1846	CTGCCCCCTGCTCACTCTCTCTGTGGGGAGAGTGAAGTGTGAGGAAGTGTGCGAG	1905
OY	621	HisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgArgCysArgAsp	640
Db	1906	CACGTCAGAGCTCATCCAGACAGTGGGGAGACAGCGTGTGGGAAGGGCGTCCCGAATC	1965
OY	641	LeuLeuLeuGlnLeuTrpLeuGlnArgProGluLeuArgValProValProGluValLeu	660
Db	1966	CTTCTCTCGACGCTCTTCAACAGCGCGGACCTCGGGGTGCCGTGCTGAGGTCTCTA	2025
OY	661	LeuHisSerGlnGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuIleHisArg	680
Db	2026	CTGCACACCGAAGGGCTGCCACAGCAGCGTGTGAACTGTGACGAGCTATCTCACCCG	2085
OY	681	PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnLysAsnArgGlyAlaAsp	700
Db	2086	TTTCAATCAAGCTCTTTCGGGACACACAGCACTCCCGGGCTTGGAAACACGAGGGCGAT	2145
OY	701	AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuLeuAlaHisLeu	720
Db	2146	GCCAGCATGGCTTCCGGAAAGCTGGGGGTGGCCACCCCGCTCTCTCTGTCGGGACCTG	2205
OY	721	ProMetIleAlaAlaLeuLeuHisGlyArgTrpHisLeuAsnPheGlnGlnPheArgGln	740
Db	2206	CCCATTGATCGGGGCTCTGCAACGGCGGACCACTCAATCTCCAGAGATTCGGGACG	2265
OY	741	GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGlnProHis	760
Db	2266	CAGAACCACTGACACTCTTCTGTCACAGTGTGGGCTCTGTGGACTGTGTGAGCGCGAC	2325

QY	761	Vahp	hkr	ser	gln	hsl	ngl	val	aleu	trp	asp	cy	leu	leu	ser	phe	ile	arg	leu	780				
Db	2326	GTGTTCCGAC	CGACGAC	CAACAGG	GGGGGG	GGCTGG	GAGATG	CTCTCT	CTTCA	TCCGGCTG														
QY	781	Leu	leu	asn	trp	arg	lys	ser	ser	arg	his	leu	ala	ala	phe	ile	asn	lys	phe	val	gln			
Db	2386	CTGCTG	AATTAC	AGAGAG	TGCTCTCC	CGCAG	CTCTG	CGCTT	CATCA	CACAG	TTGTG	CAG												
QY	801	Phe	ile	his	lys	trp	ile	thr	trp	asn	ala	pro	ala	ala	ile	ser	phe	leu	gln	his	820			
Db	2446	TTTCA	TTCATT	AGATG	ACTTAC	CTCAAT	ATGCCCC	CGACG	CCATCT	CTCTCC	TTCG	CACAG	CAC											
QY	821	Ala	asp	pro	leu	his	asp	leu	ser	phe	asp	asn	ser	asp	leu	val	met	leu	lys	ser	leu	840		
Db	2506	GCCGAC	CCGGCC	CAACG	ACGCTG	CTCTTC	GCAGAA	CAAG	AGAC	CTGTG	TATG	CTGAA	AAT	CCCTC										
QY	841	Leu	ala	gln	lys	ser	leu	pro	ser	arg	asp	asp	arg	trp	his	asp	arg	gln	leu	asn	pro	gln	860	
Db	2566	CTTGAG	AGGGCT	CACCC	CTGCCC	CACAG	GAGAC	AGACAG	AGAC	CCGAC	CGAGCC	CTG	SAC	CAAG										
QY	861	Gly	glu	gln	gln	ser	ser	ala	gln	ser	leu	pro	leu	val	ser	val	ser	leu	phe	thr	pro	880		
Db	2626	GCCGAG	GAGAG	AGACT	CAGCCG	GGCTCT	CTTGG	CCCTG	TGACG	CTCTCC	CTGTT	CAC	CCCT											
QY	881	Leu	thr	ala	ala	gln	met	ala	pro	trp	met	lys	arg	leu	ser	arg	gln	thr	val	gln	900			
Db	2686	CTGAC	CCGGG	CGGCG	GAAGG	GGCCCC	CTCAT	ATAA	AGGCTT	CCCGGG	CCAA	CGGTG	GAG											
QY	901	Asp	leu	leu	gln	gln	val	leu	ser	asp	ile	asp	gln	met	ser	arg	arg	arg	pro	gln	leu	920		
Db	2746	GATCTG	CGGAG	GGTTCT	GAGTGA	CATAG	ACAG	AGATGT	CCCCGG	CGAG	ACCCG	AGAC	CGA	AGAT	CTTG									
QY	921	Ser	phe	phe	ser	thr	asn	leu	gln	arg	leu	met	ser	ser	ala	gln	gln	uc	ys	arg	asn	940		
Db	2806	AGCTT	CTCTCT	CGACCA	CACTG	ACAGCG	GGCTG	ATGAG	CTCGG	CGAG	AGAG	AGTGG	TCGCG	CAAC										
QY	941	Leu	ala	phe	ser	leu	ala	leu	ala	arg	ser	met	gln	asn	ser	pro	ser	trp	ile	ala	ala	phe	960	
Db	2866	CTCGCT	TCAC	CCCGG	CGGCTCG	CGCTCAT	GTCA	GAAC	AGCC	CCGAG	ATTG	CAC	CCCTT	TC										
QY	961	Leu	pro	thr	phe	met	trp	cy	lys	leu	gln	ser	gln	asp	phe	gln	val	val	gln	thr	ala	leu	980	
Db	2926	CTGCC	CACTT	CATG	ATCTACT	GCTG	CTTG	GGCAG	CCAG	AGACTT	TGAG	TGTG	TCAG	AC	GGCCCTC									
QY	981	Arg	asn	leu	pro	gln	trp	ala	leu	leu	cy	gln	gln	ile	his	ala	ala	ala	val	leu	leu	his	arg	1000
Db	2986	CGGAA	CCGCTG	AGTGA	GTAGG	CTCTCT	GTG	CCAA	GAGAC	AGCGG	CTGTG	CTG	CTC	CA	CGG									
QY	1001	Ala	phe	leu	val	gln	met	trp	gln	met	asp	pro	ser	ala	gln	ile	ser	gln	ala	leu			1020	
Db	3046	GCCTT	CCGCTG	GGGCA	GTAGT	ACGG	CCAA	ANTG	AGAC	CCCG	CGCAG	ATCT	CTCC	GAG	CCCTG									
QY	1021	Arg	ile	leu	his	met	gln	ala	val	met													1029	
Db	3106	AGGAT	CTCTG	ATGAT	GAGG	CGCTG	ATG																	

APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P2930R1C8
 CURRENT APPLICATION NUMBER: US/10/033,326
 CURRENT FILING DATE: 2001-12-27
 PRIOR APPLICATION NUMBER: 60/095,325
 PRIOR FILING DATE: 1998-08-04
 PRIOR APPLICATION NUMBER: 60/112,851
 PRIOR FILING DATE: 1998-12-16
 PRIOR APPLICATION NUMBER: 60/113,145
 PRIOR FILING DATE: 1998-12-16
 PRIOR APPLICATION NUMBER: 60/113,511
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/115,558
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115,565
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115,733
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/119,341
 PRIOR FILING DATE: 1999-02-09
 PRIOR APPLICATION NUMBER: 60/119,537
 PRIOR FILING DATE: 1999-02-10
 PRIOR APPLICATION NUMBER: 60/119,965
 PRIOR FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: 60/162,506
 PRIOR FILING DATE: 1999-10-29
 PRIOR APPLICATION NUMBER: 60/170,262
 PRIOR FILING DATE: 1999-12-09
 PRIOR APPLICATION NUMBER: 60/187,202
 PRIOR FILING DATE: 2000-03-03
 PRIOR APPLICATION NUMBER: PCT/US99/12252
 PRIOR FILING DATE: 1999-06-02
 PRIOR APPLICATION NUMBER: PCT/US99/28634
 PRIOR FILING DATE: 1999-12-01
 PRIOR APPLICATION NUMBER: PCT/US99/28551
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US00/03565
 PRIOR FILING DATE: 2000-02-11
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: PCT/US00/05841
 PRIOR FILING DATE: 2000-03-02
 PRIOR APPLICATION NUMBER: PCT/US00/08439
 PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: PCT/US00/14941
 PRIOR FILING DATE: 2000-05-30
 PRIOR APPLICATION NUMBER: PCT/US00/15264
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: 2000-12-01
 NUMBER OF SEQ ID NOS: 38
 SEQ ID NO 21
 LENGTH: 3437
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-033-326-21

Alignment Scores:
 Pred. No.: 0 Length: 3437
 Score: 5241.00 Matches: 1029
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-929-769-7 (1-1029) x US-10-033-326-21 (1-3437)
 Qy 1 MethSILeuleuValaHlSalameValaHleuleuThrLeuGlyProPArgala 20
 Db 46 ARGCAATCTGCTGCTCATGCGCATGATGATCTGCTGAGCGTGGCGCCGCTCGAGCC 105

Qy 21 AspAspSerGluPheGlnAlaLeuAspIleTrpPheProGluGlyProLeuPro 40
 Db 106 GAGCAGACGAGATGCTCCAGCGCTGCTGACATGCTGCTCCGAGAGACCACTGCC 165
 Qy 41 ThrAlaPheLeuValAspThrSerGluGluAlaLeuLeuProAspTrpLeuVal 60
 Db 166 ACCGCTCTGCTGGAGACATGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTG 225
 Qy 61 ArgMetIleArgSerGluValaLeuArgLeuValaAspAlaAlaLeuGlnAspLeuGluPro 80
 Db 226 CGCATGATCGCTTGTAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
 Qy 81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyTyrLeuProValSerMetSerLeu 100
 Db 286 CAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
 Qy 101 LeuGlnPheLeuAspGlnAlaValaAlaAspProGlnThrLeuGlnGlnAsnIleMet 120
 Db 346 CTCACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 405
 Qy 121 AspLysAsnTyrMetAlaHlSalameValaGlnGlnIleGlnValaArgGlyAlaSerGly 140
 Db 406 GACAGAAATTACATGAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 465
 Qy 141 GlnThrPheHlSerLeuLeuThrAlaSerLeuProArgArgAspSerThrGluAla 160
 Db 466 CAGACTTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
 Qy 161 ProLysProLysSerSerProGlnGlnProIleGlnGlnIleArgValaGlyThr 180
 Db 526 CCCAAGCAATGAGCAGCCAGAGCAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAG 585
 Qy 181 GlnLeuArgValaLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro 200
 Db 586 CACCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
 Qy 201 LeuSerProAspProArgTrpGlnSerSerSerProArgProValAlaLeuAlaLeuGln 220
 Db 646 CTCAGCCCGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
 Qy 221 GlnAlaLeuGlnGlnGlnLeuAlaArgValaGlnGlnIleGlnSerProGlnValaProGlyIle 240
 Db 706 CAGGCCCTGGCCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
 Qy 241 ThrValArgValaLeuGlnAlaLeuAlaThrLeuLeuSerSerProHlGlyAlaLeu 260
 Db 766 ACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
 Qy 261 ValMetSerMetHlSalameValaHlSalameValaHlSalameValaHlSalameVala 280
 Db 826 GTGATGCTCATGACCGTACCGTACCGTACCGTACCGTACCGTACCGTACCGTAC 885
 Qy 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeuLeu 300
 Db 886 TACCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
 Qy 301 GlnMetLeuGlnTrpLeuAspSerProGlyValaGlnGlyProLeuArgAlaGlnLeu 320
 Db 946 CAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
 Qy 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValaArgGlyGlyLeu 340
 Db 1006 AGGATGCTTGCAGCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
 Qy 341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGluValaSerSerThrVal 360
 Db 1066 CTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
 Qy 361 ArgAlaValaIleAlaThrLeuArgSerGlyGlnGlnCysSerValaGlnProAspLeuIle 380
 Db 1126 CGTGGCTGATCGCCAGCCCTGAGAGTGTGGGAGACAGTGCAGTGCAGCCGAGCTGATC 1185
 Qy 381 SerLysValaLeuGlnGlnIleGlnValaArgSerProHlSerLeuGlnIleLeuLeuThr 400

Db	1186	AGCAAACTCTCCAGGGGGCTGATCCAGAGTGAGTCCCCCACTTGAGAGAGCTGTA	1245
Oy	401	AlaPhePheSerAlaThrAlaAspAlaAlaSerProPheProAlaCysLysProVal	420
Db	1246	GCATTCTTCTGTGCACATCGGGATGCTGCCTCCCGATTCCAGCTGTAAAGCCATT	1305
Oy	421	ValValSerSerLeuLeuLeuGlnGlnGlnProLeuAlaGlyValLysProGlyAla	440
Db	1306	GTGGTGAGCTCCCTGCTGTGCAGAGAGAGAGAGAGCCCTGGCTGGGGGGAAC	1365
Oy	441	AspGlyLysSerLeuGlnAlaValAlaGlyGlyProSerSerGlyLeuLeuValAsp	460
Db	1366	GAGGGTGGACGCTGAGAGCGCTGGGGCTGGGCCCTCTCAGAGGCTCTTAGTGA	1425
Oy	461	LeuGlnMetLeuAspProGlnValValSerSerCysProAspLeuGlnLeuArg	480
Db	1426	CTGGAATAGCTGTGACCCCGAGGTGTACACAGCTGCCCGACCTCAGCTAGG	1485
Oy	481	PheSerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTyrLeu	500
Db	1486	TTTCTCCGGAGAGAGGGCAAGGTGTACAGCCAGAGTCCCTCTTCCTGCTTAC	1545
Oy	501	ThrLeuPheThrHisGlnSerSerTyrProThrLeuHisGlnCysIleArgValLeu	520
Db	1546	ACCCCTTTCACGATCAGTCCACCTGAGCCCACTGCACAGTGCATCGAGTCTGT	1605
Oy	521	GlyLysSerArgGlnGlnArgPheAspProSerAlaSerLeuAspPheLeuTyrAla	540
Db	1606	GGCAAGAGCCGGAGAACAGAGTTGCAGCCCTGTCTGTGACTTCTGTGGCC	1665
Oy	541	IleHisValProArgIleTyrGlnGlnArgAspGlnArgThrProGlnLysArgArg	560
Db	1666	ATCATGTCTCTGCATCTGGCAGAGGGGGGAGACAGCGACCCCGCAGAACG	1725
Oy	561	GluLeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGluLeuAla	580
Db	1726	GAGTGTGTGTGGGGGTCCAGGGCCCGAGCTCATCAGCTGTGGAGCTGATCT	1785
Oy	581	GluValGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAla	600
Db	1786	GAGCGGAGAGCCGAGACCAGAGCGGGGACACAGCGCTGACGCTTATCCAGG	1845
Oy	601	LeuProLeuLeuLeuSerCysCysGlyAspAspGluSerValArgLysValThrG	620
Db	1846	CTGCGCTGTGCTCAGACTCTGCTGTGGGGAGAGATGAGAGTGTGAGAGGT	1905
Oy	621	HisLeuSerGlyCysIleGlnGlnIleTyrPylAspSerValLeuGlyIleArgC	640
Db	1906	CACCTGTAGGGCTGTACATCCAGAGTGGGGAGAGCGTGTGGAAAGGCGTCC	1965
Oy	641	LeuLeuLeuGlnLeuTyrLeuGlnArgProGluLeuArgValProValProGln	660
Db	1966	CTTCTCTGACGCTTACCTACAGGGCCGGAGCTGGGGTCCCGTCTGAGAGT	2025
Oy	661	LeuHisSerGlnGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuIleHis	680
Db	2026	CTGACACGCAAGGGCGTGCACACACAGCTGTGCATGACCTGAGAGACTAT	2085
Oy	681	PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnIleAspGlyAla	700
Db	2086	TTTATCATCGCTCTGTGGGACACACAGAGTCCCGGGCTTGGAAACGAGGGG	2145
Oy	701	AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuArgHisLeu	720
Db	2146	GCCAGCATGGCTTCGCCGAGAGCTGGCGGTGGGCAACCCCTCTCTCAGAG	2205
Oy	721	ProMetIleAlaAlaLeuLeuHisGlyArgThrHisLeuAspPheGlnIlePheArg	740
Db	2206	CCCATGATCGGGGGCTCTGTGCAGGGCCGACCCACTAATCTCAGAGAGTCC	2265
Oy	741	GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGln	760

Db	2266	CAGAACCAACCTGGACCTCTTCCTCGACAGTCTGGGCTCTGTGAGCTGTGTGACGGCCGAC	23226
Qy	761	ValPheargSerGluHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleargLeu	780
Db	2236	GTGTTCCTCCACGAGACACAGGGGGGCGTGTGGAGCTGGCTTCCTGCTTCATCCGGCTG	2385
Qy	781	LeuLeuasnTrpArgIysSerSerArgHisIleuAlaPheIleasnIysPheValGln	800
Db	2386	CTGGCAATTACAGAGAAAGTCTCCCGGCATCTGGCTGCTTCATCAACAAGTTTGTGAG	2445
Qy	801	PheIleHisIysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnIysHis	820
Db	2446	TTCATCCATAGTAGTACATTACTACAAATAGCCCCAGACGACCATCTCCCTCTGCAAGAGAC	2505
Qy	821	AlaaspProLeuHisaspLeuSerPheaspAsnSeraspLeuValIleLeuIysSerIleu	840
Db	2506	GCCGACCGCGTCCACGACCTGCTCTTCGACACAGACTGGTGTGTGTAATCCCTTC	2565
Qy	841	LeuAlaGlyLeuSerIleuProSerArgaspAspArgTrpAspArgIyLeuaspGlnGlu	860
Db	2566	CTTGCAAGGGCTTCACCTCTGCCCCAGAGGAGACAGACAGACCGACCGAGCTTGACCAAG	2625
Qy	861	GlyGlnGlnIuSerSerAspIagIySerLeuProLeuValSerValSerLeuPheThrPro	880
Db	2626	GCCGAGGAGAGACTCAGCCGCGCTCTTCCCTGGTACAGCTGCTGTTACCCCT	2685
Qy	881	LeuThrAlaAlaGluMetAlaProTyrMetIlyArgLeuSerArgIyGlnThrValGlu	900
Db	2686	CTGACCGGGCGCGAGATGGCCCCCTCATATAAACGGCTTTCCGGGGCCAAACGGTGGAG	2745
Qy	901	AspLeuLeuGluValIleuSeraspIleaspGluMetSerArgArgTrpProGlnIleLeu	920
Db	2746	GATCGCTGGAGGTTCGTAGTAGCATAGACAGATGCTCCCGGGGAGACCCGAGATCTGTG	2805
Qy	921	SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGlnIuCySerArgasn	940
Db	2806	AGCTTCCTTCGACCAACCTGCACAGCGGCTGATAGCTGGGCGAGGAGTGGTGGCCGAC	2865
Qy	941	LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaIleAlaPhe	960
Db	2866	CTCGGCTTCACACCTGGGCGCTGGCTGCATGCACAAACAGCCCGCATTTGCACGCCCTTC	2925
Qy	961	LeuProThrPheMetTrpCysLeuGlySerGlnAspPheGluValGlnThrAlaLeu	980
Db	2926	CTGCCACAGTTCATGTACTGCTGTGGGACGACAGCATTTGAGGTGGTGCAGACGGCCCTC	2985
Qy	981	ArgasnLeuProGluTyrAlaIleLeuCysGlnGluHisAlaAlaValLeuLeuHisArg	1000
Db	2986	CGGAACCTGCTGAGTACGCTCTCTCTGTGCCAAGACAGCGGCTGTGCTCCACCGG	3045
Qy	1001	AlaPheLeuValGlyMetTrpGlyGlnMetaspProSerAlaGlnIleSerGluAlaLeu	1020
Db	3046	GCCCTTCGTGGGCGCATGTACGGCCAGATGAGACCCAGCGGCGAGATCTCGAGGCCCTG	3105
Qy	1021	ArgIleLeuHisMetGluAlaValMet	1029
Db	3106	AGGATCTCGATATGAGAGCGCGTGATG	3132
RESULT 13			
US-09-822-849A-276			
Sequence 276, Application US/09822849A			
Patent No. US20020045170A1			
GENERAL INFORMATION:			
APPLICANT: Wong, Gordon G.			
APPLICANT: Clark, Hilary			
APPLICANT: Fechtel, Kim			
APPLICANT: Agostino, Michael J.			
APPLICANT: Howes, Steven H.			
APPLICANT: Resnick, Richard J.			
APPLICANT: Gulukota, Kamalakara			
APPLICANT: Graham, James R.			
APPLICANT: Genetics Institute, Inc.			
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS			

RESULT 13
 US-09-822-849A-276
 ; Sequence 276, Application US/09822849A
 ; Patent No. US20020045170A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mong, Gordon G.
 ; APPLICANT: Clark, Hilary
 ; APPLICANT: Fechtel, Kim
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Howes, Steven H.
 ; APPLICANT: Resnick, Richard J.
 ; APPLICANT: Gulukota, Kamalakara
 ; APPLICANT: Graham, James R.
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS

```

: FILE REFERENCE: GIN 6403
: CURRENT APPLICATION NUMBER: US/09/822,849A
: CURRENT FILING DATE: 2001-09-04
: PRIOR APPLICATION NUMBER: 60/195,582
: PRIOR FILING DATE: 2000-04-06
: NUMBER OF SEQ ID NOS: 598
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 276
: LENGTH: 2824
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-822-849A-276

Alignment Scores:
Pred. No.: 0          Length: 2824
Score: 4211.00       Matches: 829
Percent Similarity: 98.81%  Conservative: 1
Best Local Similarity: 98.69%  Mismatches: 0
Query Match: 80.35%  Indels: 10
DB: 10              Gaps: 1

US-09-929-769-7 (1-1029) x US-09-822-849A-276 (1-2824)

QY 200 ProLeuSerProaspProargTrpGlnSerSerSerProargProValAlaLeuAlaLeu 219
Db 2 CCGCTCAGCCCGGACCTCGGTGGCAGAGCTCCAGCCCGCGGCTCGCTGCGCTG 61
QY 220 GlnGlnAlaLeuGlnGlnGlnLeuAlaArgValAlaGlnGlnGlnGlnGlnGln 239
Db 62 CACAGAGCCCTGGGCGCAGAGCTGGCCGCGCTCCAGGGGACCCCGAGTCCGGCC 121
QY 240 IleThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyAla 259
Db 122 ATCAGCGTGGTGTCTCCAGAGCCCTGGCCACCTCTCAGCTCCACACAGCGCGTGC 181
QY 260 LeuValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeu 279
Db 182 CTGGTGTGTGCATGACCGCTAGCCACTCTGCGCGCGCTGCGCGCTGCGCGCTGCG 241
QY 280 GlnTyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeu 299
Db 242 CAGTACAGAGCGGTGTGGCAGACAGGAGCGGTCTCTCGCTCTCTCTCTCTCTCTC 301
QY 300 LeuGlnMetLeuGlnTrpLeuAspSerProGlyValGlnGlnGlnGlnGlnGln 319
Db 302 CTCAGATGCTGAGTGGCTGGACAGCCCTGGCGTGGAGGGCGCGCTGCGGGCACAG 361
QY 320 LeuArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGly 339
Db 362 CTCAGGATGCTTGGCAGCCAGGCGCTCAGCGGGGCGAGGCTCACTGATGTGCGAGGG 421
QY 340 LeuLeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGlnValAlaSerSerThr 359
Db 422 CTCCTGGCGCTGGCGGCGCGCTGCGCTTCCGTAGAGCTGAGAGTGTCCAGCTCAC 481
QY 360 ValArgAlaValAlaLeuAlaThrLeuArgSerGlyGlnGlnGlnGlnGlnGln 379
Db 482 GTCCGTCCTCTCATCGCCAGCCCTGAGGTCTGGGGACAGTGCAGCTGGAGCCGAGCTG 541
QY 380 IleSerIysValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 399
Db 542 ATCAGCAAACTCTCCAGGGGGCTGATCGAGGTGAGGTCCCGCCAGCTGGAGAGAGCTCTG 601
QY 400 ThrAlaPhePheSerAlaThrAlaAspAlaAlaSerProPheProAlaCysIysProVal 419
Db 602 ACTGCAATCTTCTCTGCACTGGGAGTGTGCTCCCGCTTCCAGCTGTAAAGCCGCTT 661
QY 420 ValValAlaSerSerLeuLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 439
Db 662 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 721
QY 440 AlaAspGlyIysLeuGlnGlnAlaValArgLeuGlyProSerSerGlyLeuLeuValAsp 459

```

```

Db 722 GCGAGCGGTGGCAGCCCTGAGAGCCGTGCGGCGCTGCGCTGCTCAGTGGAC 781
QY 460 TrpLeuGlnMetLeuAspProGlnValAlaSerSerCysProAspLeuGlnLeuArgLeu 479
Db 782 TGGCTGGAATGCTGGAGCCCGGAGGTGTGACAGAGCTGCCCGACCTCGAGCTAGGCTG 841
QY 480 LeuPheSerArgArgGlyGlyGlyGlnAlaGlnAlaProSerPheArgProTyrLeu 499
Db 842 CTCCTTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 901
QY 500 LeuThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnGlnIleArgValLeu 519
Db 902 CTGACCTCTTCCAGCATCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 961
QY 520 LeuGlyIysSerArgGlnGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAla 539
Db 962 CTGGGCAAGAGCGGGAGAACAGAGGTTCGACCCCTGCTCTGAGACTTCTCTGGGCC 1021
QY 540 CysIleHisValProArgIleTrpGlnGlyArgAspGlnArgThrProGlnIysArgArg 559
Db 1022 TGCATCCATGTCCTCCATCTGCGCAGGGCGGGAGCCAGCCCGCAGAGAGCGCGCG 1081
QY 560 GlnGlnLeuValLeuArgValGlnGlyProGlnLeuLeuIleSerLeuValGlnLeu 579
Db 1082 GAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1141
QY 580 AlaGlnAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAla 599
Db 1142 GCGAGGCGGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1201
QY 600 ArgLeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgIysValThr 619
Db 1202 CGCGTCCCGCTGCTGCTCACTGCTGTGGGACATGAGATGTCCAGAGAGCTACAG 1261
QY 620 GlnHisLeuSerGlyCysIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 639
Db 1262 GAGCAGCTGTGAGCTCATTCACAGTGGGAGACAGCTGCTGGCGAGCGCTGCCGA 1321
QY 640 AspLeuLeuGlnGlnLeuTyrLeuGlnArgProGlnLeuArgValProValProGlnVal 659
Db 1322 GACCTTCTCGTGCAGCTCTACCTACAGCGGCGGAGCTGGCGGCTGCTGAGAGTCT 1381
QY 660 LeuLeuHisSerGlnGlyAlaAlaSerSerSerValCysIysLeuAspGlyLeuIleHis 679
Db 1382 CTACTGCACAGCAAGGCGGTGCCAGCAGCAGCTCTGCAAGCTGAGCGGCTCATATCAC 1441
QY 680 ArgPheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnValAspArgGlyAla 699
Db 1442 CGCTTCATCAGCTCTTCCGAGACCAAGCAGTCCCGGGCGTTGGAGAACCGAGGGCG 1501
QY 700 AspAlaSerMetAlaCysArgIysLeuAlaValAlaHisProLeuLeuLeuAlaArgHis 719
Db 1502 GATGCCAGCATAGGCTCGCGGAAGCTGGCGGTGGCGGACCGGCTGTGCTCAGGCAC 1561
QY 720 LeuPrometIleAlaAlaLeuLeuHisGlyArgThrHisLeuAsnPheGlnGlnPheArg 739
Db 1562 CTGCCCATGATCGGGGCGCTCTGACAGGGCGGACCCAGCTCACTCCAGGAGATTCGG 1621
QY 740 GlnGlnAsnHisLeuSerCysPheLeuHisValLeuGlnGlnLeuGlnGlnPro 759
Db 1622 CAGCAGAACCACTCAGAGCTGCTCTCAGCTGTGGGCTGCTGAGAGTCTCTACACCG 1681
QY 760 HisValPheArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArg 779
Db 1682 CAGCTGTTCGAGCGAGCAGCAGAGGGGCGCTGGGAGTGTGTGTCTTCATATCCGC 1741
QY 780 LeuLeuLeuAsnTyrArgIysSerSerArgHisLeuAlaAlaPheIleAsnIysPheVal 799
Db 1742 CTGCTGTGAATTAACAGGAAGTCTCCCGCATCTGCTGCTCAACCAAGATTGTG 1801
QY 800 GlnPheIleHisIysTyrIleThrTyrAsnAlaProAlaAlaIleSerPheLeuGlnIys 819
Db 1802 CAGTTCATCCATAGTACATTACCTACATGCCCCAGCAGCCATCTCTCTGCGAGAG 1861

```

Oy	820	HlsAlaAspProLeuHisAspIleuSerPheAspAsnSerAspLeuValMetIleuLysSer	833
Db	1862	CACCCGACCCGGCTCCACGACCTGTCTTCGACAAACAGGACCTGGTATGTGTAAATCC	1921
Oy	840	LeuLeuAlaGlyLeuSerIleuProSerArgAspAspArgThrAspArgIYleuAspSlu	859
Db	1922	CTCTTCGAGGGCTCACCCTCGCCACAGGAGACAGACAGACCAGAGCTTGACGAA	1961
Oy	860	GluGlyGluGluGluSerSerAspIleuGlySerLeuProLeuValSerValSerLeuPheThr	879
Db	1982	GAGGGCAGAGAGGAGACCTCAGCCGGCTCTTCGCCCTGGTACAGGCTCTCCCTGTTACAC	2041
Oy	880	ProLeuThrAlaAlaGluMetAlaProTyrMetLysArgLeuSerArgGlyGlnThrVal	899
Db	2042	CTCTTGACCCGGGGCCGAGATGGCCCCCTACATGAAACGGCTTTCCGGGGCCAAACGGGTG	2101
Oy	900	Glu-----AspLeuLeuGluValLeuSerAspIle	909
Db	2102	GAGGGTGAGTACAGGCCCTGCTTACCCACGACGACAGATCTCTGGAGTTCGTAGTACATA	2161
Oy	910	AspGluMetSerArgArgArgProGluIleLeuSerPhePheSerThrAsnLeuGlnArg	929
Db	2162	GACAGAGTGTCCCGCGCGAGAACCCGAGATCCGTGACTTCTTCTCGACCAACCTGCAGCGG	2221
Oy	930	LeuMetSerSerAlaGluGlyCysArgAsnLeuAlaPheSerLeuAlaLeuArgSer	949
Db	2222	CTGATGAGCTGGCGCGAGGAGTGTTCGCGAACCTTCGCTTACGCTGGCCCTGGCGTCC	2281
Oy	950	MetGlnAsnSerProSerIleAlaAlaAlaPheLeuProThrPheMetTyrCysLeuGly	969
Db	2282	ATGCAGAAACAGCCCCACACTTGGACGGCGTTCTCTCCACGTTGATGACGCTTGCGC	2341
Oy	970	SerGlnAspPheGluValValGlnThrAlaLeuArgAsnLeuProGluTyrAlaLeuLeu	989
Db	2342	AGCCAGAGACTTGGAGTGGTGGCAGAGCGCCCTCCGGAACCTCCATGAGTACGCTCTCTG	2401
Oy	990	CysGlnGlnHisAlaAlaValIleuLeuHisArgAlaPheLeuValGlyMetTyrGlyGln	1009
Db	2402	TGCCAAGACGACGGCGGCTGTGCTGCTCCACCGGGCTTCTCTGTGGGCACTGTGACGGCCAG	2461
Oy	1010	MetAspProSerAlaGlnIleSerGluAlaLeuArgIleLeuHisMetGluAlaValMet	1029
Db	2462	ATGACCCCCAGCGGCGAGATCTCCGAGGCCCTGAGGATCTGCATATGAGAGCGGTGATG	2521
RESULT 14			
US-09-984-271-56			
: Sequence 56, Application US/09984271			
: Publication No. US20030040088A1			
: GENERAL INFORMATION:			
: APPLICANT: Rosen et al.			
: TITLE OF INVENTION: 71 Human Secreted Proteins			
: FILE REFERENCE: P2030P1			
: CURRENT APPLICATION NUMBER: US/09/984, 271			
: CURRENT FILING DATE: 2001-10-29			
: PRIOR APPLICATION NUMBER: 09/482, 273			
: PRIOR FILING DATE: 2000-01-13			
: PRIOR APPLICATION NUMBER: PCT/US99/15849			
: PRIOR FILING DATE: 1999-07-14			
: PRIOR APPLICATION NUMBER: 60/092, 921			
: PRIOR FILING DATE: 1998-07-15			
: PRIOR APPLICATION NUMBER: 60/092, 922			
: PRIOR FILING DATE: 1998-07-15			
: PRIOR APPLICATION NUMBER: 60/092, 956			
: PRIOR FILING DATE: 1998-07-15			
: NUMBER OF SEQ ID NOS: 267			
: SOFTWARE: PatentIn Ver. 2.0			
: SEQ ID NO 56			
: LENGTH: 1559			
: TYPE: DNA			
: ORGANISM: Homo sapiens			
: FEATURE:			
: NAME/KEY: SITE			

```

; LOCATION: (1445)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1551)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-984-271-56

Alignment Scores:
Pred. No.: 6,566-212 Length: 1559
Score: 2055.00 Matches: 404
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.21% Indels: 0
DB: 9 Gaps: 0
US-09-929-769-7 (1-1029) x US-09-984-271-56 (1-1550)

```

QY 906 LeuSerAspIleAspGluMetSerArgArgProGluIleLeuSerPhePheSerThr 925
 DB 841 CTGAGTACCTATAGACGAGTGTCCCGCGGAGACCCGAGATCTGACCTTCTCTGACC 900
 QY 926 AsnLeuGlnArgLeuMetSerSerAlaGluGluCysCysArgAsnLeuAlaPheSerLeu 945
 DB 901 AACCTGACGGGGTGTATGAGCTGCGCGAGAGATGTTGCGCAACCTCGCTTCAGCTTG 960
 QY 946 AlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaAlaPheLeuProThrPheMet 965
 DB 961 GCCCTCGCTCCATCAGAAACACCCGAGATGTCAGACCCCTTCCGCGGAGCTTCATG 1020
 QY 966 TyrCysLeuGlySerGlnSerPheGluValAlaGlnThrAlaLeuArgAsnLeuProGlu 965
 DB 1021 TACTGCTGGGCGACGAGCTTTGAGTGTGTCAGACGCGCCCTCGGAACCTGCTGAG 1080
 QY 986 TyrAlaLeuLeuGlyGlnGlnIleAlaAlaValLeuLeuHisArgAlaPheLeuValGly 1005
 DB 1081 TACGCTCTCTGTGCGCAAGACGACGCGCTGTCTGCTCCACCGGCTTCCTGCTGAGG 1140
 QY 1006 MetTyrGlyGlnMetAspProSerAlaGlnIleSerGluAlaLeuArgIleLeuHisMet 1025
 DB 1141 ATGTAGCGCCAGATGAGACCCGAGCGGAGATCTCCGAGGCCCTGAGAGATCTGCATATG 1200
 QY 1026 GluAlaValMet 1029
 DB 1201 GAGCGCGTGTATG 1212

RESULT 15
 US-10-040-739-1131/c
 ; Sequence 1131: Application US/10040739
 ; Patent No. US2002017635A1

GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 McCoy, John
 Lavalley, Edward
 Racie, Lisa
 Merberg, David
 Treacy, Maurice
 Spaulding, Vikki

TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
 NUMBER OF SEQUENCES: 1519

CORRESPONDENCE ADDRESS:
 ADDRESSER: Genetics Institute, Inc.
 STREET: 87 Cambridgepark Drive
 CITY: Cambridge
 STATE: Massachusetts
 COUNTRY: U.S.A
 ZIP: 02140

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/040,739
 FILING DATE: 07-Jan-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION NUMBER: 09/036,520
 FILING DATE: 03-JUN-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Brown, Scott A.
 REGISTRATION NUMBER: 32,724
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 1131:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 639 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 1131:
 US-10-040-739-1131

Alignment Scores:
 Pred. No.: 2 62e-90 Length: 639
 Score: 924.00 Matches: 188
 Percent Similarity: 98.44% Conservative: 1
 Best Local Similarity: 97.92% Mismatches: 2
 Query Match: 17.63% Indels: 1
 DB: 9 Gaps: 0

US-09-929-769-7 (1-1029) x US-10-040-739-1131 (1-639)

QY 839 SerLeuLeuAlaGlyLeuSerLeuProSerArgAspArgThrAspArgGlyLeuAsp 858
 DB 637 TCCTCTCTGACGAGGCTCAGACCTGCCAGAGGAGGACGACGACGACGAGGCTGAG 578
 QY 859 GluGluGlyGluGluGluSerSerAlaGlySerLeuProLeuValSerValSerLeuPhe 878
 DB 577 GAAGAGGCGAGGAGGAGAGACTCAGCGGCTCTGCGGCTGAGCGTCTCCCTGTTG 518
 QY 879 ThrProLeuThrAlaAlaGluMetAlaProTyrMetLysArgLeuSerArgGlyGlnThr 898
 DB 517 ACCCTCTGACCGCGCGGCGAGATGCGCCCTACATGAAACGGCTTTCGCGGCGCAACG 458
 QY 899 ValGluAspLeuGluGluValLeuSerAspIleAspGluMetSerArgArgArgProGlu 918
 DB 457 GTGAGAGATGCTGTGAGGTTCTGAGTATGACGAGATGTCGCGGCGGAGACCGAG 398
 QY 919 IleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGluGlyCys 938
 DB 397 ATCCTGAGCTTCTTCTCGAACCACTGACGCGGCTGATGAGTGGCGCGGAGGTGTTC 338
 QY 939 ArgAsnLeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAla 958
 DB 337 CGCAACCTGCGCTTCAAGCCCTGCGCTCCATGACAGACAGCCCGCNCATTCAGACC 278
 QY 959 AlaPheLeu-ProThrPheMetTyrCysLeuGlySerGlnAspPheGluValAlaGlnThr 978
 DB 277 GCTTCTGCGCCNACGTTCTATGACTGCGCGGCGACGAGAGATGAGTGTGCGAGAC 218
 QY 978 IAlaLeuArgAsnLeuProGluTyrAlaLeuGlyGlnGlnIleAlaAlaValLeu 998
 DB 217 GGCCTTCGGAACCTGCTGATGAGCTCTCTGTCACAGACAGCGGCTGTCTGCT 158
 QY 998 uHisArgAlaPheLeuValGlyMetTyrGlyGlnMetAspProSerAlaGlnIleSerG 1018
 DB 157 CCACGCGGCTTCTCTGTGTGGCATGTACGGCCAGATGAGACCCGAGCGGAGATCTCCGA 98
 QY 1018 uAlaLeuArgIleLeuHisMetGluAlaValMet 1029
 DB 97 GGCCTGAGAGATCTCATATGAGGCGCGTAGG 64

Search completed: May 14, 2003, 02:38:44
 Job time : 442 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:41:38 ; Search time 44 Seconds
(without alignments)
3116.247 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 1029
Sequence: 1 MHILVHAMVILLTLPRA.....MDPSAQISEALRIILMEAVM 1029

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: A_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/genescp-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/genescp-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/genescp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1029	100.0	1029	21 AAB24027	Human PRO3434 prot
2	1029	100.0	1029	21 AAY96736	PRO3434, a novel s
3	1029	100.0	1029	23 AAU07765	Tumour associated
4	504	49.0	650	21 AAB42495	Human ORFX ORF2259
5	404	39.3	404	21 AAY87196	Human secreted pro
6	404	39.3	404	22 AAE06173	Human gene 46 enco
7	404	39.3	404	23 ABG33997	Human secreted pro
8	327	31.8	327	22 AAE06086	Human gene 46 enco
9	327	31.8	327	23 ABG33908	Human secreted pro
10	327	31.8	328	21 AAY87109	Human secreted pro

11	43	4.2	161	22 AAM95497	Human reproductive
12	8	0.8	55	22 AAU52278	Protonibacterium
13	8	0.8	84	22 AAU64883	Protonibacterium
14	8	0.8	101	23 ABP34288	Human kinase-like
15	8	0.8	207	22 ABB58449	Drosophila melanog
16	8	0.8	209	23 AAU83137	Novel secreted pro
17	8	0.8	217	22 AAC81704	S. epidermidis ope
18	8	0.8	224	20 AAW72977	Rhodobacter sphaer
19	8	0.8	224	20 AAW76766	R. sphaeroides Adh
20	8	0.8	226	22 AAB2495	Formaldehyde trans
21	8	0.8	253	20 AAY08016	Mouse PA-I protein
22	8	0.8	311	22 ABB59518	Drosophila melanog
23	8	0.8	390	20 AAY05302	S. aureus protein
24	8	0.8	396	23 ABB39735	Staphylococcus epi
25	8	0.8	705	22 AAB94141	Human protein sequ
26	8	0.8	853	22 AAB95413	Human protein sequ
27	8	0.8	893	22 AAB93648	Human protein sequ
28	8	0.8	1137	22 AAB94750	Human protein sequ
29	8	0.8	1846	22 AAU00295	Interferon induced
30	8	0.8	2042	22 ABB59689	Drosophila melanog
31	7	0.7	27	18 AAM09063	Epsitein-Barr virus
32	7	0.7	27	18 AAM18617	Peptide #5051 enco
33	7	0.7	27	22 AAM31077	Peptide #5114 enco
34	7	0.7	27	23 ABB40522	Human peptide enco
35	7	0.7	35	23 AAU10698	cPLA2-adaptor pept
36	7	0.7	36	9 AAB80687	Human secreted pro
37	7	0.7	38	22 ABB50760	Human secreted pro
38	7	0.7	39	21 AAB63080	Human secreted pro
39	7	0.7	43	21 AAG55810	Arabidopsis thalia
40	7	0.7	46	21 AAG59085	Gene 4, human secre
41	7	0.7	47	21 AAB34147	Plant SP encoded
42	7	0.7	47	21 AAB25068	Human secreted alp
43	7	0.7	48	22 AAB60198	Novel central nerv
44	7	0.7	48	22 AAB87142	Human ORFX protein
45	7	0.7	50	23 ABB00236	

ALIGNMENTS

RESULT 1	
AAB24027	standard: Protein: 1029 AA.
ID	AAB24027
AC	AAB24027
DT	25-JAN-2001 (first entry)
DE	Human PRO3434 protein sequence SEQ ID NO:23.
KW	Human; tumour; diagnosis; neoplastic disease; proliferation; cancer; Identification; tumorigenesis; anticancer; detection.
OS	Homo sapiens.
PN	WO200053750-A1.
PD	14-SEP-2000.
PF	02-DEC-1999; 99WO-US28551.
PR	01-SEP-1999; 99WO-US05028.
PR	29-OCT-1999; 99US-0162506.
PR	30-NOV-1999; 99WO-US28313.
PR	01-DEC-1999; 99WO-US28634.
PA	(GETH) GENENTECH INC.
PI	Botstein D, Goddard A, Gurney AL, Roy MA, Matanabe CX, Wood WT.
DR	WPI: 2000-594320/56.
DR	N-PSDB: AAC58109.

XX Antibodies specific for PRO polypeptides, used to diagnose and inhibit
 PT the growth of tumors in mammals, and to identify inhibitors of PRO
 PT polypeptide activity or expression -
 XX
 PS Claim 61: Fig 14; 226pp; English.
 XX
 CC The present invention describes an antibody that binds to a human
 CC protein (I) selected from: PRO381; PRO1269; PRO1410; PRO1755; PRO1780;
 CC PRO3434; PRO1927; PRO3555; PRO1096; PRO2038; and PRO2262. (I) has
 CC anticancer activity and can be used to diagnose tumors in mammals, by
 CC detecting complex formation when the antibody is contacted with test
 CC cells. Increased expression of genes encoding (I) can also be detected
 CC to diagnose tumors. Agents which inhibit the activity of (I),
 CC especially the antibodies, or an antisense oligonucleotide which
 CC hybridises to genes encoding (I), can be used to inhibit tumour growth,
 CC preferably by inducing cell death. Methods from the present invention
 CC can be used to identify compounds which inhibit the biological activity
 CC of (I). AAC58019 to AAC58102 represent PCR primers and hybridisation
 CC probes used in examples from the present invention for human PRO
 CC sequences. AAC58103 to AAC58122 and AAB24021 to AAB24040 represent human
 CC PRO polynucleotide and protein sequences given in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 1029 AA:
 Query Match 100.0%; Score 1029; DB 21; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 601 LPILLSCCGDDSVKRVTEHLSGCIQWGDVYGRRCRDLLOLYQRELPVPEVL 660
 Qy 661 LHSEGAASSVCKIDGILHFRITLLADTSRALENAGDASMACRKLAVANPILLRL 720
 Db 661 LHSEGAASSVCKIDGILHFRITLLADTSRALENAGDASMACRKLAVANPILLRL 720
 Qy 721 PMIAALLHGRTHLNFQFROQNLSCFLHVLGLLELLQPHVFRSEHOGALMDCLLSTRL 780
 Db 721 PMIAALLHGRTHLNFQFROQNLSCFLHVLGLLELLQPHVFRSEHOGALMDCLLSTRL 780
 Qy 781 LWNKRSRLAFLINKEVQFIHKYTYNAPPAISFLQKHADPLHDLSFNSDLVYKSL 840
 Db 781 LWNKRSRLAFLINKEVQFIHKYTYNAPPAISFLQKHADPLHDLSFNSDLVYKSL 840
 Qy 841 LAGSLPSRDRDRDRCDEGEESAGSLPVSVSLFTPLTAENAPYKRLSRGQTV 900
 Db 841 LAGSLPSRDRDRDRCDEGEESAGSLPVSVSLFTPLTAENAPYKRLSRGQTV 900
 Qy 901 DLLEVLSDIDEMSRRELLSFSTNLQRLMSSAECCRNLAESLALRSMQNSPSIAAF 960
 Db 901 DLLEVLSDIDEMSRRELLSFSTNLQRLMSSAECCRNLAESLALRSMQNSPSIAAF 960
 Qy 961 LPTFMVCLGSQDEEVQVOTALRNLPEYALLCOEHAVALHRAFLVGMVGMDPSAQISEAL 1020
 Db 961 LPTFMVCLGSQDEEVQVOTALRNLPEYALLCOEHAVALHRAFLVGMVGMDPSAQISEAL 1020
 Qy 1021 RILHMEAVM 1029
 Db 1021 RILHMEAVM 1029

RESULT 2
 AAY96736
 ID AAY96736 standard; Protein: 1029 AA.
 AC AAY96736;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE PRO3434, a novel secreted protein.
 XX
 KW PRO3434; secreted protein; transmembrane protein; recombinant production;
 KM gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide 1..16
 FT /label= Signal_peptide
 FT Modified-site 91..97
 FT /note= "N-myristoylation site"
 FT Modified-site 136..142
 FT /note= "N-myristoylation site"
 FT Modified-site 154..158
 FT /note= "CAMP- and cGMP-dependent protein kinase
 FT phosphorylation site"
 FT Modified-site 224..230
 FT /note= "N-myristoylation site"
 FT Modified-site 329..333
 FT /note= "amidation site"
 FT Modified-site 331..335
 FT /note= "CAMP- and cGMP-dependent protein kinase
 FT phosphorylation site"
 FT Modified-site 435..441
 FT /note= "N-myristoylation site"
 FT Modified-site 439..445
 FT /note= "N-myristoylation site"
 FT Modified-site 443..449
 FT /note= "N-myristoylation site"
 FT Modified-site 616..620
 FT /note= "CAMP- and cGMP-dependent protein kinase
 FT phosphorylation site"
 FT Modified-site 634..638

```

FT      /note= "amidation site"
FT      665..671
FT      /note= "N-myristoylation site"
FT      698..704
FT      /note= "N-myristoylation site"
FT      783..789
FT      /note= "cAMP- and cGMP-dependent protein kinase
FT      phosphorylation site"
FT      891..895
FT      /note= "cAMP- and cGMP-dependent protein kinase
FT      phosphorylation site"
XX      WO200036102-A2.
XX      22-JUN-2000.
XX      01-DEC-1999; 99WO-US28634.
XX      16-DEC-1998; 98US-0112851.
XX      16-DEC-1998; 98US-0113145.
XX      22-DEC-1998; 98US-0113511.
XX      12-JAN-1999; 99US-0115558.
XX      12-JAN-1999; 99US-0115565.
XX      12-JAN-1999; 99US-0115733.
XX      09-FEB-1999; 99US-0119341.
XX      10-FEB-1999; 99US-0119537.
XX      12-FEB-1999; 99US-0119665.
XX      02-JUN-1999; 99WO-US12252.
XX      (GETH ) GENENTECH INC.
XX      Botstein D, Desnovers L, Ferrara N, Fong S, Gao W, Goddard A;
XX      Guney AL, Pan J, Roy MA, Stewart TA, Tumas D, Watanabe CK;
XX      Wood WT;
XX      WPI: 2000-431586/37.
XX      N-PSDB: AAA51266.
XX      Isolated nucleic acid molecule encodes a PRO polypeptide which is a
XX      transmembrane polypeptide
XX      Claim 12; Fig 16; 154pp: English.
XX      This is PRO3434, a novel secreted protein. The invention
XX      concerns novel secreted and transmembrane proteins, designated PRO
XX      polypeptides. The cDNA and gene sequences are useful in the recombinant
XX      production of PRO polypeptides, as a hybridization probe to screen
XX      libraries to isolate cDNAs with sequence identity to PRO polypeptides or
XX      to map the gene encoding the PRO polypeptides and analyzing genetic
XX      disorders. The cDNA/gene can also be used to produce transgenic animals
XX      useful for the development and screening of therapeutically useful
XX      reagents. They can also be used in gene therapy, e.g. to replace a
XX      defective gene.
XX      Sequence 1029 AA:
XX      Query Match 100.0%; Score 1029; DB 21; Length 1029;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB      181 QLRVLGPEDDLQAGMFLQIFPLSPDRMOSSSPRPVALALQALGQELARVVOGSFEVPQI 240
QY      241 TVRVLQALATLLSPHGALVMSMRSHFLACPLRLQCLQYRCYRPODTGESSLELKVL 300
DB      241 TVRVLQALATLLSPHGALVMSMRSHFLACPLRLQCLQYRCYRPODTGESSLELKVL 300
QY      301 QMLQWLDSPGVGCGPLRAQRLMLASOASAGRLSDVRGGLRLAEALAFRODLEVVSTV 360
DB      301 QMLQWLDSPGVGCGPLRAQRLMLASOASAGRLSDVRGGLRLAEALAFRODLEVVSTV 360
QY      361 RAVIATLRSGECCSVPEPDLISKVLOGLIEVRSRPHLEELLTAFFSATADAASFPACKPVY 420
DB      361 RAVIATLRSGECCSVPEPDLISKVLOGLIEVRSRPHLEELLTAFFSATADAASFPACKPVY 420
QY      421 VVSSILLQEEEPRLAGCKRGADGSLAVRLGSSGLVDMLEMDPEVYSSCPDQLRL 480
DB      421 VVSSILLQEEEPRLAGCKRGADGSLAVRLGSSGLVDMLEMDPEVYSSCPDQLRL 480
QY      481 FSRRKKGQAGVPSFRPYLLFTTHQSSWPTLHQCIRVLLGKSRQRPDPASLDPLMAC 540
DB      481 FSRRKKGQAGVPSFRPYLLFTTHQSSWPTLHQCIRVLLGKSRQRPDPASLDPLMAC 540
QY      541 IHVPRIWGRDQRTPOKRREELVLRVQGPBELISVELLAEAEFRSODGDTAAGSLIQAR 600
DB      541 IHVPRIWGRDQRTPOKRREELVLRVQGPBELISVELLAEAEFRSODGDTAAGSLIQAR 600
QY      601 LPLLISCCCGDESVRKVTEHLSCIQMGDSVLGRCHDLLOLYLORPEELRVVPEVL 660
DB      601 LPLLISCCCGDESVRKVTEHLSCIQMGDSVLGRCHDLLOLYLORPEELRVVPEVL 660
QY      661 LHSEGAASSVCKLQGLIHREFTLLADTSDSRALENRGADASMACRKLAVAHPLLLRL 720
DB      661 LHSEGAASSVCKLQGLIHREFTLLADTSDSRALENRGADASMACRKLAVAHPLLLRL 720
QY      721 PMIAALLHGRTHLNFQEFROQNHLSCLFLHVLGLLELLQPHVFRSEHOGALMCCLSFIRL 780
DB      721 PMIAALLHGRTHLNFQEFROQNHLSCLFLHVLGLLELLQPHVFRSEHOGALMCCLSFIRL 780
QY      781 LLYNRKSSRHLLAFINKEFQFIHKYITTYNAPPAISFLQAHAPLDLSDNSDIWLKSL 840
DB      781 LLYNRKSSRHLLAFINKEFQFIHKYITTYNAPPAISFLQAHAPLDLSDNSDIWLKSL 840
QY      841 LAGLSLPSHDRTDGLGEEGEESSAGSLPVYSVLFPTLTAEMAPYMKRLSRQOTVE 900
DB      841 LAGLSLPSHDRTDGLGEEGEESSAGSLPVYSVLFPTLTAEMAPYMKRLSRQOTVE 900
QY      901 DLLEVLSDIDEMSRRPETLSPFTNLQRLMSABECCRNLAFLSALRSMQSPSTIAAF 960
DB      901 DLLEVLSDIDEMSRRPETLSPFTNLQRLMSABECCRNLAFLSALRSMQSPSTIAAF 960
QY      961 LPTFMVCGSDQFEVYQVTRNLPEYALICQHAVALHRAFLVGMYGMDPSAQISEAL 1020
DB      961 LPTFMVCGSDQFEVYQVTRNLPEYALICQHAVALHRAFLVGMYGMDPSAQISEAL 1020
QY      1021 RLHMEAVM 1029
DB      1021 RLHMEAVM 1029
XX      RESULT 3
XX      ID AU077765
XX      AU077765 standard; Protein; 1029 AA.
XX      05-JUN-2002 (first entry)
XX      Tumour associated antigenic target polypeptide (TAT) 156.
XX      Tumour associated antigenic target polypeptide: TAT; cancer:
XX      breast cancer; colorectal cancer; lung cancer; ovarian cancer;
XX      central nervous system cancer; liver cancer; bladder cancer;
XX      KM

```

KW	pancreatic cancer; cervical cancer; melanoma; leukaemia; TAI156.
XX	
OS	Homo sapiens.
XX	
XX	Key
FT	Location/Qualifiers
FT	1..16
FT	/label= Signal_peptide
FT	17..1029
FT	/label= Mature_TAI156
FT	/note= "Tumour associated antigenic target polypeptide"
FT	91..97
FT	/label= N-myristoylation_site
FT	136..142
FT	/label= N-myristoylation_site
FT	154..158
FT	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT	224..230
FT	/label= N-myristoylation_site
FT	329..333
FT	/label= Amidation_site
FT	331..335
FT	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT	435..441
FT	/label= N-myristoylation_site
FT	439..445
FT	/label= N-myristoylation_site
FT	443..449
FT	/label= N-myristoylation_site
FT	616..620
FT	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT	634..638
FT	/label= Amidation_site
FT	665..671
FT	/label= N-myristoylation_site
FT	698..704
FT	/label= N-myristoylation_site
FT	785..789
FT	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT	891..895
FT	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT	
XX	WO200216581-A2.
XX	
PD	28-FEB-2002.
XX	
PF	14-AUG-2001: 2001WO-US25464.
XX	
PR	24-AUG-2000: 2000WO-US23328.
PR	28-FEB-2001: 2001WO-US06520.
PR	22-JUN-2001: 2001US-0888257.
PR	22-JUN-2001: 2001WO-US20118.
XX	
PA	(GENTECH) GENENTECH INC.
XX	
PI	Gao W, Polakis P, Shou J, Smith V, Soriano R, Williams PM, Wu TD;
PI	Zhang Z;
XX	
DR	WPI: 2002-280928/32.
XX	
XX	N-PSDB; ABK11743.
XX	
PT	Novel isolated antibody which binds to tumour-associated antigenic target polypeptide useful for killing cancer cells expressing the polypeptide and for treating tumour comprising cells that expresses the
XX	
PS	Claim 2; Fig 7; 123pp; English.
XX	

CC The invention describes an isolated antibody which binds to a
 CC tumour-associated antigenic target (TAT) polypeptide. The antibody is
 CC useful for: killing a cancer cell (such as a breast, colorectal, lung,
 CC ovarian, central nervous system, liver, bladder, pancreatic, cervical,
 CC melanoma or leukaemia cell) that expresses a polypeptide with at least
 CC 80% identity to the TAT polypeptide sequence; treating a tumour
 CC comprising cells that express a polypeptide with at least 80% identity
 CC to the TAT polypeptide sequence; determining the presence of a
 CC polypeptide having at least 80% identity to the TAT polypeptide sequence
 CC in a sample suspected of containing the polypeptide; diagnosing the
 CC presence of a tumour in a mammal, and for antibody dependent enzyme
 CC mediated prodrug therapy (ADEPT). This is the amino acid sequence of
 CC the tumour associated antigenic target polypeptide (TAT) 156, described
 CC in the invention.

SO Sequence 1029 AA;

Query Match 100.0%; Score 1029; DB 23; Length 1029;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1029; Conservative 0;

QY	1	MHLVYHAWYILTLGPPRADSEFQALDIWPEEKPLPTATLVPTSEALLPDWLKL	60
DB	1	MHLVYHAWYILTLGPPRADSEFQALDIWPEEKPLPTATLVPTSEALLPDWLKL	60
QY	61	RMIRSEVLRLVDAALQDLPEQQLLFVQSGFIPVSSMSKLLQFLDQAVAHDPQTEONIM	120
DB	61	RMIRSEVLRLVDAALQDLPEQQLLFVQSGFIPVSSMSKLLQFLDQAVAHDPQTEONIM	120
QY	121	DKNYMAHVEVQHERGASGGQTFHSLTSLAPRRRSTAPKPKSSPEQITGGRLRVGT	180
DB	121	DKNYMAHVEVQHERGASGGQTFHSLTSLAPRRRSTAPKPKSSPEQITGGRLRVGT	180
QY	181	QLRVLPEDDLAMFQIIFPLSPDPMWSSPPVALALQALQGLARVQSSPEVPGI	240
DB	181	QLRVLPEDDLAMFQIIFPLSPDPMWSSPPVALALQALQGLARVQSSPEVPGI	240
QY	241	TVRVQALATLTLSSPHGALVMSMRSHFLACPLRQLCQYQRCVQDTGFSSLFKVL	300
DB	241	TVRVQALATLTLSSPHGALVMSMRSHFLACPLRQLCQYQRCVQDTGFSSLFKVL	300
QY	301	OMLQWLDSPGVEGGLRAQLRMLASQASAGRRLSDVRGGILRLAEALAFRODLEVVSTV	360
DB	301	OMLQWLDSPGVEGGLRAQLRMLASQASAGRRLSDVRGGILRLAEALAFRODLEVVSTV	360
QY	361	RAVIATLRSGEQSVSPDLISKVLQGLIEVRSPHEBELLTAFPSATADASPAPACKPVY	420
DB	361	RAVIATLRSGEQSVSPDLISKVLQGLIEVRSPHEBELLTAFPSATADASPAPACKPVY	420
QY	421	VVSSLLOEEEPPLAGKPGADGSLFAVRLGESSGLLVMLDPEVYSSCPDQLRL	480
DB	421	VVSSLLOEEEPPLAGKPGADGSLFAVRLGESSGLLVMLDPEVYSSCPDQLRL	480
QY	481	FSRRKRGQAQVPSFRPYLLTFTHOSWPTLHQICRVLVLGSRREDFPSASLDLMLC	540
DB	481	FSRRKRGQAQVPSFRPYLLTFTHOSWPTLHQICRVLVLGSRREDFPSASLDLMLC	540
QY	541	IHWPRITMGSDQRTPOKRREBELVRYQGPGLISVLLEAEATRSQDDDTACSLIQAR	600
DB	541	IHWPRITMGSDQRTPOKRREBELVRYQGPGLISVLLEAEATRSQDDDTACSLIQAR	600
QY	601	LPILLSCCGGDESVKRVTEHLISGCIQWGDVYLGRCDLQLQLVQREPLRVPEVL	660
DB	601	LPILLSCCGGDESVKRVTEHLISGCIQWGDVYLGRCDLQLQLVQREPLRVPEVL	660
QY	661	LHSEGAASSVCKLDGLIHREITLADTSDSRALREMGADASACKLVAHPDLLRL	720
DB	661	LHSEGAASSVCKLDGLIHREITLADTSDSRALREMGADASACKLVAHPDLLRL	720
QY	721	PMIAALHGRTHLNFQEFQONHLSCLVHLGLLELLQHVRSERHQGLAMDCILSFIRL	780
DB	721	PMIAALHGRTHLNFQEFQONHLSCLVHLGLLELLQHVRSERHQGLAMDCILSFIRL	780

QY 761 LNNRKSRRHLAAFINKEVQFIHKYITYNAPAAISFLQKHADPLHLSFDNSDLVMLKSL 840
DB 761 LNNRKSRRHLAAFINKEVQFIHKYITYNAPAAISFLQKHADPLHLSFDNSDLVMLKSL 840
QY 841 LAGSLPRDRDRTDGLDEEGEESASGLPLVSYSLFPLTAAMAPYMKRLSGQTYE 900
DB 841 LAGSLPRDRDRTDGLDEEGEESASGLPLVSYSLFPLTAAMAPYMKRLSGQTYE 900
QY 901 DLLEVLSDIDEMSRRRPELISFSTNLRQIMSSAECCRNLAFLSLMSQSPSTAAAF 960
DB 901 DLLEVLSDIDEMSRRRPELISFSTNLRQIMSSAECCRNLAFLSLMSQSPSTAAAF 960
QY 961 LPTFWYCLGSDQFEVYQALRNLPYVALLCOEHAVALHRAFLVGMQGMPSAQLISEAL 1020
DB 961 LPTFWYCLGSDQFEVYQALRNLPYVALLCOEHAVALHRAFLVGMQGMPSAQLISEAL 1020
QY 1021 RLHMEAVM 1029
DB 1021 RLHMEAVM 1029

RESULT 4
ID AAB42495 standard; Protein: 650 AA.
AC AAB42495;
XX 08-FEB-2001 (first entry)
DE Human ORFX ORF2259 polypeptide sequence SEQ ID NO:4518.
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
KW vulnereary; antiparkinsonian; antiparkinsonian; neuroprotective;
KW anticonvulsant; osteopathic; antirheumatic; immunosuppressant; cardiatic;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antineoplastic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX Homo sapiens.
OS MO200058473-A2.
PN 05-OCT-2000.
PD 31-MAR-2000; 2000MO-US08621.
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
PA Shinkets RA, Leach M;
PI WPI: 2000-602362/57.
DR N-PSDB: AAC76704.
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
PS Claim 11, Page 3717-3718; 5507pp; English.
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytosolic; hepatotropic; vulnereary;
CC antiparkinsonian; antiparkinsonian; neuroprotective;
CC osteopathic; anticonvulsant; antirheumatic; immunosuppressant;
CC immunostimulant; cardiatic; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antineoplastic; antibacterial; antifungal; antineoplastic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

Sequence 650 AA:

Query Match 49.0%; Score 504; DB 21; Length 650;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 397 ELTFAFFSATADAPFPACKRVVYVSSLLDDEEPLAGKRGAGGSLAENRLGPSGL 456
DB 8 ELTFAFFSATADAPFPACKRVVYVSSLLDDEEPLAGKRGAGGSLAENRLGPSGL 67
QY 457 LVDWLEMLDPEVYSSCPDLQLRLFSRRKGGGOAOPSPRPULLTLFTHSSMPTLHOCI 516
DB 68 LVDWLEMLDPEVYSSCPDLQLRLFSRRKGGGOAOPSPRPULLTLFTHSSMPTLHOCI 127
QY 517 RVLGKSRQRPDPASLDFLMACIHVPRIWGRDQRTPOKREELVLRVQPELISLVE 576
DB 128 RVLGKSRQRPDPASLDFLMACIHVPRIWGRDQRTPOKREELVLRVQPELISLVE 187
QY 577 LTLAEFRSQGDPACSLIOARLPULLSCCGDDESVKRYTEHLSCGTOQWGSVYLR 636
DB 188 LTLAEFRSQGDPACSLIOARLPULLSCCGDDESVKRYTEHLSCGTOQWGSVYLR 247
QY 637 RCRDLLQLYLRPELRVVPVLLHSEGAASSVCKLDGLHRTTLADTSDSRALEN 696
DB 248 RCRDLLQLYLRPELRVVPVLLHSEGAASSVCKLDGLHRTTLADTSDSRALEN 307
QY 697 RGADSMACRKLAVAHPLLLRHLPIMIALHGRTHLNFQEFROQNHLSCTLHVGLLEL 756
DB 308 RGADSMACRKLAVAHPLLLRHLPIMIALHGRTHLNFQEFROQNHLSCTLHVGLLEL 367
QY 757 LQPHVFRSEHOGALWDCLSLFRLILNKRKSRHLAAFINKEVQFIHKYITYNAPAAISF 816
DB 368 LQPHVFRSEHOGALWDCLSLFRLILNKRKSRHLAAFINKEVQFIHKYITYNAPAAISF 427
QY 817 LQKHADPLHLSFDNSDLVMLKSLAGLSLPRDRDRTDGLDEEGEESASGLPLVSYSL 876
DB 428 LQKHADPLHLSFDNSDLVMLKSLAGLSLPRDRDRTDGLDEEGEESASGLPLVSYSL 487
QY 877 LFTPLTAAMAPYMKRLSGQTYE 900
DB 488 LFTPLTAAMAPYMKRLSGQTYE 511
RESULT 5
ID AAY87196 standard; Protein: 404 AA.
AC AAY87196;
XX 09-MAY-2000 (first entry)
DE Human secreted protein sequence SEQ ID NO:235.

XX Human: secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
 KW tumour; neurodegenerative disorder; developmental abnormality; allergy;
 KW foetal deficiency; blood disorder; immune system disorder; arthritis;
 KW autoimmune disease; hepatic disease; renal disease; inflammation;
 KW Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
 KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 KW asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 KW reproductive disorder; gastrointestinal disorder; respiratory disorder;
 KW metabolic disorder; food additive; preservative.
 XX Homo sapiens.
 OS WO200004140-A1.
 PN 27-JAN-2000.
 PD 14-JUL-1999; 99WO-US15849.
 PF 15-JUL-1998; 98US-0092921.
 PR 15-JUL-1998; 98US-0092922.
 PR 15-JUL-1998; 98US-0092956.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
 PI Mucenski M, Edress GA, Soppet DR;
 DR WPI: 2000-161128/14.
 DR N-PSDB: AA298062.
 XX New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers, neurological or blood disorders
 PT Disclosure: page 466-467, 494pp; English.
 PS The polynucleotide sequences given in AA298017 to AA298108 encode the
 CC human secreted proteins given in AA87064 to AA87223. Human secreted
 CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Human secreted proteins and their polynucleotides can
 CC be used for developing products for the diagnosis or treatment of cancer,
 CC tumours, neurodegenerative disorders, developmental abnormalities and
 CC foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal diseases, inflammation,
 CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The
 CC proteins or polynucleotides can also be used as food additives or
 CC preservatives. The proteins are also useful for identifying their
 CC binding partners. AA298008 to AA298016 and AA87063 are sequence used in
 CC the exemplification of the present invention.
 XX Sequence 404 AA:
 SQ
 Query Match 39.3%; Score 404; DB 21; Length 404;
 Best Local Similarity 100.0%; Pred No. 0; Mismatches 0; Indels 0; Caps 0;
 Matches 404; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 QY 626 IQQMGDSVLRRCRDLLQLYLRPELRVPEVLLHSEGAASSVCKDLGILHRTTL 685
 DB 1 IQQMGDSVLRRCRDLLQLYLRPELRVPEVLLHSEGAASSVCKDLGILHRTTL 60

QY 686 ADTSRRALENRGADASMACRKLAVAHPIILLRLHPMIAALLHGRTHLNFQEFROQNHLS 745
 DB 61 ADTSRRALENRGADASMACRKLAVAHPIILLRLHPMIAALLHGRTHLNFQEFROQNHLS 120
 QY 746 CFHIVGLLELLQPHVFRSHQALNDCLLSFIRLLNLNKKSRHLAFLNKKVQFIHKY 805
 DB 121 CFHIVGLLELLQPHVFRSHQALNDCLLSFIRLLNLNKKSRHLAFLNKKVQFIHKY 180
 QY 806 ITYNAPAIIFLQKHADPLHDLSFDNSDVLMLKSLAGLSLPRDRDTRDGLDEEGEES 865
 DB 181 ITYNAPAIIFLQKHADPLHDLSFDNSDVLMLKSLAGLSLPRDRDTRDGLDEEGEES 240
 QY 866 SAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRRPEILSFEST 925
 DB 241 SAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRRPEILSFEST 300
 QY 926 NIGRLSSAECCRNLAFLSALRSMONSPSIAAFLPTFMVCLGSODFEVQATRLRLPE 985
 DB 301 NIGRLSSAECCRNLAFLSALRSMONSPSIAAFLPTFMVCLGSODFEVQATRLRLPE 360
 QY 986 YALLCGEHAIVLLHRAFLVGMGOMPSSAQISEALRLHMEAYM 1029
 DB 361 YALLCGEHAIVLLHRAFLVGMGOMPSSAQISEALRLHMEAYM 404

RESULT 6
 AA06173
 ID AA06173 standard; Protein; 404 AA.
 XX AA06173;
 DT 24-SEP-2001 (first entry)
 XX Human gene 46 encoded secreted protein fragment, SEQ ID NO:235.
 DE Human: secreted protein; proliferative disorder; cancer; tumour; asthma;
 XX foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KW cell culture; chemotaxis; vulnerability; binding partner identification;
 KW gene therapy.
 XX Homo sapiens.
 OS WO200151504-A1.
 PN 19-JUL-2001.
 PD 12-JAN-2001; 2001WO-US00911.
 PF 13-JAN-2000; 2000US-0482273.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Edress GA, Mucenski M, Ebner R;
 DR WPI: 2001-425865/45.
 DR Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition
 PT Disclosure: page 116; 864pp; English.
 PS AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
 CC protein genes, and AA06041-AA06132 represent the proteins they encode.
 CC AA06133-AA06205 represent human secreted protein fragments.
 CC The secreted proteins and their genes are useful for preventing, treating

CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 71 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC The present sequence represents a human secreted protein fragment
CC referred to in the disclosure of the invention.

CC
XX Sequence 404 AA:

Query Match 39.3%; Score 404; DB 22; Length 404;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 100MGDSVIGRCRDLLOLYLQRPRLRPVEVLLHSGAASSVCKIDGLHRRITLL 685
DB 1 100MGDSVIGRCRDLLOLYLQRPRLRPVEVLLHSGAASSVCKIDGLHRRITLL 60
OY 686 ADTSRRALENGADASMACRKLAVAHPLLLRHLPMTAALHGRTHLNFQFROONHLS 745
DB 61 ADTSRRALENGADASMACRKLAVAHPLLLRHLPMTAALHGRTHLNFQFROONHLS 120
OY 746 CETHLVGLLELLQPHVFRSEHOGALMDCILSTRILLNTRKSSRHAAINKEFVQFIHXY 805
DB 121 CETHLVGLLELLQPHVFRSEHOGALMDCILSTRILLNTRKSSRHAAINKEFVQFIHXY 180
OY 806 IYYNPAAPISFQKHADPLHDSFNSDPLVMLKSLAGLSLSPDRDTRRGDEEEERS 865
DB 181 IYYNPAAPISFQKHADPLHDSFNSDPLVMLKSLAGLSLSPDRDTRRGDEEEERS 240
OY 866 SAGSLPLVSVSLFTPLTAEMAPYMKRSLRGQTVEDLEVLSDIDEMSRREPILTSFEST 925
DB 241 SAGSLPLVSVSLFTPLTAEMAPYMKRSLRGQTVEDLEVLSDIDEMSRREPILTSFEST 300
OY 926 NIQRLMSSAECCRNLAFLARSMQNSPISIAAEPLFTMYCLGSDFEVOTALRNLP 985
DB 301 NIQRLMSSAECCRNLAFLARSMQNSPISIAAEPLFTMYCLGSDFEVOTALRNLP 360
OY 986 YALLCOEHAVALHRAFLVGMGTGMDPSQOISEALRIIMEAVM 1029
DB 361 YALLCOEHAVALHRAFLVGMGTGMDPSQOISEALRIIMEAVM 404

RESULT 7
ABG33997
ID ABG33997 standard; Protein; 404 AA.
XX
AC ABG33997;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human secreted protein #2 encoded by gene 46.
XX
KW Human; secreted protein; gene therapy; immunosuppressive;
KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;

KW vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
KW virucide; fungicide; ophthalmological; autoimmune disease; neoplasm;
KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; skin aging;
KW epithelial cell proliferation; food additive.

OS Homo sapiens.

PN WO200226931-A2.

PD 04-APR-2002.

PE 24-SEP-2001; 2001WO-US29871.

PR 25-SEP-2000; 2000US-234925P.

PR 12-JAN-2001; 2001WO-US00911.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;
PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
PI Endress GA, Mucenski M, Ebner R;

PI MPI: 2002-362489/39.

DR N-PSDB: ABK69771.

XX Novel 71 isolated secreted polypeptides and polynucleotides encoding
PT the polypeptides, useful for treating Huntington's disease, sepsis,
PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis,
PT asthma

PS Disclosure: Page 1313-1314; 1478pp; English.

CC The invention relates to an isolated nucleic acid molecule (or its
CC fragment, homologue complement or allelic variant) encoding a human
CC secreted protein (and its fragment, domain, epitope, variant, secreted
CC form and species variant). Also included are a recombinant vector
CC comprising the nucleic acid, a recombinant host cell comprising the
CC vector, an antibody against the secreted protein, a recombinant host cell
CC that expresses the secreted protein and a method of identifying a binding
CC partner of the secreted protein. The nucleic acid and protein are used to
CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
CC for example autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders,
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC ocular disorders e.g. corneal infection. Many other diseases and
CC disorders are listed in the specification. The polypeptides can also be
CC used to aid wound healing an epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. The present
CC sequence represents a novel human secreted protein of the invention.

XX Sequence 404 AA:

Query Match 39.3%; Score 404; DB 23; Length 404;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 100MGDSVIGRCRDLLOLYLQRPRLRPVEVLLHSGAASSVCKIDGLHRRITLL 685
DB 1 100MGDSVIGRCRDLLOLYLQRPRLRPVEVLLHSGAASSVCKIDGLHRRITLL 60
OY 686 ADTSRRALENGADASMACRKLAVAHPLLLRHLPMTAALHGRTHLNFQFROONHLS 745
DB 61 ADTSRRALENGADASMACRKLAVAHPLLLRHLPMTAALHGRTHLNFQFROONHLS 120

KM virucide; fungicide; ophthalmological; autoimmune disease; neoplasm;
 KM rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
 KM cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
 KM angiodysplasia; nervous system disorder; Alzheimer's disease; infection;
 KM ocular disorder; corneal infection; wound healing; skin aging;
 KM epithelial cell proliferation; food additive.
 OS Homo sapiens.
 XX WO200226931-A2.
 XX 04-APR-2002.
 XX 24-SEP-2001; 2001WO-US29871.
 XX 25-SEP-2000; 2000US-234925P.
 XX 12-JAN-2001; 2001WO-US00911.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Mucenski M, Edner R;
 XX WPI; 2002-362489/39.
 DR N-PSDB; AAK69771.
 XX
 PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
 PT the polypeptides, useful for treating Huntington's disease, sepsis,
 PT meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
 PT asthma
 PS Claim 11: Page 1250-1251; 1478pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (or its
 CC fragment, homologue complement or allelic variant) encoding a human
 CC secreted protein (and its fragment, domain, epitope, variant, secreted
 CC form and species variant). Also included are a recombinant vector
 CC comprising the nucleic acid, a recombinant host cell comprising the
 CC vector, an antibody against the secreted protein, a recombinant host cell
 CC that expresses the secreted protein and a method of identifying a binding
 CC partner of the secreted protein. The nucleic acid and protein are used to
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
 CC for example autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiodysplasia, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
 CC ocular disorders e.g. corneal infection. Many other diseases and
 CC disorders are listed in the specification. The polypeptides can also be
 CC used to aid wound healing an epithelial cell proliferation, to prevent
 CC skin aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities. The present
 CC sequence represents a novel human secreted protein of the invention.
 XX
 SQ Sequence 327 AA:
 Query Match 31.8%; Score 327; DB 23; Length 327;
 Best Local Similarity 100.0%; Pred. No. 1.4e-314;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 121 PLHDSFSDMSDLVMLKSLLAGLSLPSRDDRTDGLDEGEESBSAGSLPLVSVSLPTPLT 180
 QY 883 AAEAPYMRRLSGQVDELLLEVLSDIDEMSRRRPILSFPSNLRQLMSSAECCRNLA 942
 DB 181 AAEAPYMRRLSGQVDELLLEVLSDIDEMSRRRPILSFPSNLRQLMSSAECCRNLA 240
 QY 943 FSLALRSKMSNPSIAAAPTFCGSDQFEVQATALNRLPEYALCOEHAVALHRAE 1002
 DB 241 FSLALRSKMSNPSIAAAPTFCGSDQFEVQATALNRLPEYALCOEHAVALHRAE 300
 QY 1003 LVGMGQMDPSAQISEALRLHMEAVM 1029
 DB 301 LVGMGQMDPSAQISEALRLHMEAVM 327
 RESULT 10
 AA87109
 ID AA87109 standard; Protein; 328 AA.
 XX
 AC AA87109;
 DT 09-MAY-2000 (first entry)
 XX
 DE Human secreted protein sequence SEQ ID NO:148.
 XX
 KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
 KW tumour; neurodegenerative disorder; developmental abnormality; allergy;
 KW foetal deficiency; blood disorder; immune system disorder; arthritis;
 KW autoimmune disease; hepatic disease; renal disease; inflammation;
 KW Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
 KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 KW asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 KW reproductive disorder; gastrointestinal disorder; respiratory disorder;
 KW metabolic disorder; food additive; preservative.
 XX
 OS Homo sapiens.
 XX WO200004140-A1.
 XX 27-JAN-2000.
 XX 14-JUL-1999; 99WO-US15849.
 XX 15-JUL-1998; 98US-0092921.
 XX 15-JUL-1998; 98US-0092922.
 XX 15-JUL-1998; 98US-0092956.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Edner R, Olsen HS, Brewer LA, Florence KA, Young PE;
 PI Mucenski M, Endress GA, Soppet DR;
 XX WPI; 2000-161128/14.
 DR N-PSDB; AA298062.
 XX
 PT New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers, neurological or blood disorders
 XX
 CC Claim 11: Page 404-406; 494pp; English.
 CC The polynucleotide sequences given in AA298017 to AA298108 encode the
 CC human secreted proteins given in AA87064 to AA87223. Human secreted
 CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new

CC polynucleotides. Human secreted protein s and their polynucleotides can
CC be used for developing products for the diagnosis or treatment of cancer
CC tumours, neurodegenerative disorders, developmental abnormalities and
CC foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The
CC proteins or polynucleotides can also be used as food additives or
CC preservatives. The proteins are also useful for identifying their
CC binding partners. AAZ88008 to AAZ9016 and AAY87063 are sequence used in
CC the exemplification of the present invention.

SQ **Sequence** **328 AA;**

```

Query Match      31.8%  Score 327:  DB 21:  Length 328:
Best Local Similarity 100.0%:  Pred.No. 1.4e-314:
Matches 327:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0

```

DY 703 MACRLAVAHPLLLRHLPMIALDLHGRIHNLQEFERQQNHSCLFHVLGELLEQLQPHV · 76
DB 1 MACRLAVAHPLLLRHLPMIALDLHGRIHNLQEFERQQNHSCLFHVLGELLEQLQPHV · 60

0y 763 RSEHGALWOCLLSFIRLLLNTRKSSRHIAAFINKFVQFIHKYITYNAPALISFLQKHAD 82
 |||||
 Db 61 RSEHGALWOCLLSFIRLLLNTRKSSRHIAAFINKFVQFIHKYITYNAPALISFLQKHAD 120
 |||||

Db 121 PLHDSFDNSDLVMLKSLLAGLSLPSRDDRTDRGLDEEGEERSASGLPLVSVSLFTPTLT 180

003 AAAGMFMATKSRGVIYEEDLLEVLSDIDEMSKRRKEILSTFSINLQRLMSSAECCRNLA 94
07 AA
181 AAEMAPYMKRLSRGVIYEDLLEVLSDIDEMSKRRKEILSTFSINLQRLMSSAECCRNLA 24

241 FSLALRSMQNSPSIAAFLPTFWYCLGSDFEVQGTALRNPPEYALLCQEHAAVLLHRAF 300

```
Db      301 LVGMTGQMPSPAQISEALRIILHMEAVM 327
```

RESULT 11	.
AAM95497	
ID AAM95497	standard; Protein; 161 AA.

AC	AAM95497;	.
XX		
DT	21-NOV-2001 (first entry)	

DE	Human reproductive system related antigen SEQ ID NO: 4155.
XX	
KW	Human; reproductive system related antigen; reproductive system disorder; sex hormone;

XX	
XX	
OS	Homo sapiens.
XX	
PN	W0300155320-A7

AA	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001 . 2001WO-HS01339

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR

[illegible]

PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249266.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 RA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI: 2001-465570/50.
 DR N-PSDB; AAL01467.
 XX
 XX Isolated nucleic acid molecule encoding a reproductive system antigen
 PT is used in preventing, treating or ameliorating a medical condition
 XX
 XX Claim 11; SEQ ID NO 4155; 1297bp + Sequence Listing; English.
 CC
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,

CC including cancer. The present sequence is a protein of the invention.
 XX
 SO Sequence 161 AA;
 CC
 CC Query Match 4.2%; Score 43; DB 22; Length 161;
 CC Best Local Similarity 100.0%; Pred. No. 9,4e-34;
 CC Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 851 DRTDGLDEEGEESSAGSLPLVSVLFTPLTAAMAPYMKRL 893
 DB 41 DRTDGLDEEGEESSAGSLPLVSVLFTPLTAAMAPYMKRL 83
 RESULT 12
 ID AAU52278 standard; Protein; 55 AA.
 XX
 AC AAU52278;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #13174.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 RA (CORI-) CORIAX CORP.
 XX
 XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59553.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 13473; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 55 AA;
 Query Match 0.8%; Score 8; DB 22; Length 55;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 331 RRLSDVWG 338
 |||||
 Db 12 RRLSDVWG 19

RESULT 13.
 AAU64883
 ID AAU64883 standard; Protein; 84 AA.
 XX
 AC AAU64883;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #25779.
 XX
 KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN MO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001MO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 PA (CORI-) CORIXA CORP.
 PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 DR N-PSDB; AAS59654.
 DR WPI; 2001-616774/71.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 26078; 10699P; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 84 AA;
 Query Match 0.8%; Score 8; DB 22; Length 84;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 405 ATADASP 412
 |||||
 Db 3 ATADASP 10

RESULT 14
 ABP34284
 ID ABP34284 standard; Protein; 101 AA.
 XX
 AC ABP34284;
 XX
 DT 09-JUL-2002 (first entry)
 XX
 DE Human kinase-like ORF3257 protein, SEQ ID NO:6514.
 XX
 KM Human: ORF; open reading frame; ORFX; drug screening; diagnosis;
 KM disease monitoring; cytokine; cell proliferation; cell differentiation;
 KM immune modulation; haematopoiesis regulation; tissue growth;
 KM angiogenesis; actinin; inhibin; chemotactic; chemokinetic; haemostatic;
 KM thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KM behaviour; cancer; proliferative disorder; neurological disorder;
 KM cardiovascular disease; immune system disorder; organ transplantation;
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KM hypothyroidism; cholesterol ester storage disease; infection; vunerary;
 KM vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
 KM neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KM cardiant; hypotensive; antihypoid; antiinflammatory; immunomodulator;
 KM dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN MO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001MO-US17076.
 PR 24-MAY-2000; 2000US-206690P.
 PA (CURA-) CURAGEN CORP.
 PI Leach MD, Shinkets RA;
 PI N-PSDB; ABN78310.
 DR WPI; 2002-106200/14.
 DR N-PSDB; ABN78310.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 XX
 PS Claim 10; Page 1875-1876; 2508PP; English.
 XX
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,

CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokine activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antifibrotic activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 CC
 XX
 SQ Sequence 101 AA:

Query Match 0.8%; Score 8; DB 23; Length 101;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 750 VGLLELL 757
 |||||
 Db 36 VGLLELL 43

RESULT 15

ABB58449
 ID ABB58449 standard; Protein; 207 AA.

XX ABB58449;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 2139.

XX Drosophila: developmental biology; cell signalling; insecticide;
 pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL02552.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 2139; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX

SQ Sequence 207 AA:

Query Match 0.8%; Score 8; DB 22; Length 207;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 427 LOEEPEPLA 434
 |||||
 Db 145 LOEEPEPLA 152

Search completed: May 7, 2003, 19:44:31
 Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:43:23 ; Search time 27 Seconds
(without alignments)
3663.790 Million cell updates/sec

Title: US-09-929-769-7

Perfect score: 1029

Sequence: 1 MHILVHAMVILLTLCPPRA.....MDPSAQISEALRIIMEAVM 1029

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
1: PIR-73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	633	61.5	640	2 T08758	hypothetical prote
2	100	9.7	310	2 T46429	hypothetical prote
3	9	0.9	188	2 G75326	probable 16S rRNA
4	9	0.9	241	2 B75628	DNA-binding stress
5	9	0.9	327	2 A83436	hypothetical prote
6	8	0.8	157	2 T24380	hypothetical prote
7	8	0.8	272	2 E90194	hypothetical prote
8	8	0.8	279	2 G95117	Hemk protein [limpo
9	8	0.8	345	2 E88103	protein W10G11.5 [
10	8	0.8	386	2 G89801	hypothetical prote
11	8	0.8	423	2 A29393	hemocyanin beta-c
12	8	0.8	410	2 A70947	hypothetical prote
13	8	0.8	593	2 AB0046	probable N-acetyl
14	8	0.8	640	2 F75553	hypothetical prote
15	8	0.8	958	2 C87504	hypothetical prote
16	8	0.8	997	2 T32814	hypothetical prote
17	8	0.8	1154	2 A13431	hypothetical prote
18	8	0.8	1815	2 S73021	chromosome segrega
19	8	0.8	1822	2 F87203	polyketide synthas
20	7	0.7	53	2 B82583	hypothetical prote
21	7	0.7	88	2 A61356	hypothetical prote
22	7	0.7	88	2 A11726	hypothetical prote
23	7	0.7	89	2 A83465	flagellar biosynth
24	7	0.7	103	2 H71134	hypothetical prote
25	7	0.7	111	2 A63411	hypothetical cytos
26	7	0.7	113	2 A80736	probable bacteriop
27	7	0.7	128	2 E70866	probable globin -
28	7	0.7	132	2 P00060	T-cell receptor be
29	7	0.7	148	2 T37069	hypothetical prote

30	7	0.7	150	2 D69100	probable protein d
31	7	0.7	156	2 B59225	allergen Bos d 2.0
32	7	0.7	156	2 A59225	allergen Bos d 2.0
33	7	0.7	156	2 A75149	hypothetical prote
34	7	0.7	160	2 B84903	hypothetical prote
35	7	0.7	163	2 S14269	phenomene-binding
36	7	0.7	171	2 E84912	probable polyribon
37	7	0.7	183	2 F71127	hypothetical prote
38	7	0.7	185	2 A72528	hypothetical prote
39	7	0.7	188	2 T02247	hypothetical prote
40	7	0.7	194	2 T46920	hypothetical prote
41	7	0.7	196	2 E69100	hypothetical prote
42	7	0.7	197	2 F65019	hypothetical prote
43	7	0.7	198	2 B83267	probable glutathio
44	7	0.7	200	2 JN0579	DNA-binding protei
45	7	0.7	204	2 AF0914	conserved hypothet

ALIGNMENTS

RESULT 1																																							
T08758																																							
hypothetical protein DKFZP586J0619.1 - human (fragment)																																							
C:Species: Homo sapiens (man)																																							
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999																																							
C:Accession: T08758																																							
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.																																							
submitted to the Protein Sequence Database, March 1999																																							
A:Reference number: Z16471																																							
A:Accession: T08758																																							
A:Molecule type: mRNA																																							
A:Residues: 1-640 <MAN>																																							
A:Cross-references: EMBL:AL050110																																							
A:Experimental source: adult uterus; clone DKFZP586J0619																																							
C:Genetics:																																							
A:Note: DKFZP586J0619.1																																							
Query Match																																							
Best Local Similarity 61.5%; Score 633; DB 2; Length 640;																																							
Matches 633; Conservative 100.0%; Pred. No. 0;																																							
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;																																							
QY	397	ELTTAFSSATADAA	PFPA	CKPVVV	SVLL	LEEE	PLAG	KPGAD	GSLEAVRLGSSGL 456																														
DB	8	ELTTAFSSATADAA	PFPA	CKPVVV	SVLL	LEEE	PLAG	KPGAD	GSLEAVRLGSSGL 67																														
QY	457	LVDMLEMLDPEV	YSSCPD	QLRL	FLFS	RRKG	GQAQ	VPFR	PLYLTFTHQSSWPTLHOCI 516																														
DB	68	LVDMLEMLDPEV	YSSCPD	QLRL	FLFS	RRKG	GQAQ	VPFR	PLYLTFTHQSSWPTLHOCI 127																														
QY	517	RVLLGSRQRPDP	PSASL	DFLNAC	IVPR	IWGR	QDRT	POKR	REELVLRVQGPETLSVE 576																														
DB	128	RVLLGSRQRPDP	PSASL	DFLNAC	IVPR	IWGR	QDRT	POKR	REELVLRVQGPETLSVE 187																														
QY	577	LILAEETRSDGD	TAA	CSLI	QAR	PL	LLSC	CGD	DESVRKYTEHLSSGCIQMGDSVILGR 636																														
DB	188	LILAEETRSDGD	TAA	CSLI	QAR	PL	LLSC	CGD	DESVRKYTEHLSSGCIQMGDSVILGR 247																														
QY	637	KCRDLLQIYLR	PELR	RVVPE	VVLL	SEGA	SSVCK	LDGL	IHRFTTLLADTSDSALEN 696																														
DB	248	KCRDLLQIYLR	PELR	RVVPE	VVLL	SEGA	SSVCK	LDGL	IHRFTTLLADTSDSALEN 307																														
QY	697	RGADASMACR	KLAV	APHL	LLRL	PLM	IA	LLGR	TILNFQEFROQNHLSGFLHVLGILEL 756																														
DB	308	RGADASMACR	KLAV	APHL	LLRL	PLM	IA	LLGR	TILNFQEFROQNHLSGFLHVLGILEL 367																														
QY	757	LQPHVFRSEHQ	ALW	DC	LS	FTRL	LN	IVRK	SSRH	LAFTNKEVQFTHKYITYNAPAISEF 816																													
DB	368	LQPHVFRSEHQ	ALW	DC	LS	FTRL	LN	IVRK	SSRH	LAFTNKEVQFTHKYITYNAPAISEF 427																													
QY	817	LQKHADPLH	LD	IS	FD	NS	DI	VM	KS	SL	LAG	LS	LP	S	R	D	D	T	D	G	L	D	E	E	G	E	E	S	S	A	G	S	L	P	V	S	V	S	876
DB	428	LQKHADPLH	LD	IS	FD	NS	DI	VM	KS	SL	LAG	LS	LP	S	R	D	D	T	D	G	L	D	E	E	G	E	E	S	S	A	G	S	L	P	V	S	V	S	487

QY 877 LFTPLTAEMAPYMKRLSRGQTVEDLLVLSIDIDEMSRREPELISFSTNLQRLMSAAE 936
 DB 488 LFTPLTAEMAPYMKRLSRGQTVEDLLVLSIDIDEMSRREPELISFSTNLQRLMSAAE 547
 QY 937 CCNLFALFSLARFMONSPSTAAAFLEPFMYCLCSODEVQVOTLRNLPETALLCOEHAAV 996
 DB 548 CCNLFALFSLARFMONSPSTAAAFLEPFMYCLCSODEVQVOTLRNLPETALLCOEHAAV 607
 QY 997 LHRRAFLVGMVGMDSQAQISEALRLIMEAVM 1029
 DB 608 LHRRAFLVGMVGMDSQAQISEALRLIMEAVM 640
 RESULT 2
 T46429
 hypothetical protein DKFZp434C0126.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46429
 R:Ansgorge, W.; Winkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z33028
 A:Accession: T46429
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-310 <NAA>
 A:Cross-references: EMBL:AL137358
 A:Experimental source: adult testis; clone DKFZp434C0126
 C:Genetics:
 A:Note: DKFZp434C0126.1
 Query Match 9.7%; Score 9; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 3.2e-92;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 844 LSLPSRDDRDRLDIDEGEESSAGSLPLYSVSLFPTLTAEMAPYMKRLSRGQTVEDLL 903
 DB 89 LSLPSRDDRDRLDIDEGEESSAGSLPLYSVSLFPTLTAEMAPYMKRLSRGQTVEDLL 148
 QY 904 EVLSIDIDEMSRREPELISFSTNLQRLMSAAECCRNIAF 943
 DB 149 EVLSIDIDEMSRREPELISFSTNLQRLMSAAECCRNIAF 188
 RESULT 3
 G75326
 probable 16S rRNA processing protein RlmM - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: G75326
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MID:20036896; PMID:10567266
 A:Accession: G75326
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-188 <WHI>
 A:Cross-references: GB:AE002038; GB:AE000513; MID:96459790; PIDN:AAF11557.1; PID:9645979
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2010
 A:Map position: 1
 Query Match 0.9%; Score 9; DB 2; Length 188;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 432 PLAGKPGA 440
 ||||||||

DB 155 PLAGKPGA 163
 RESULT 4
 B75628
 DNA-binding stress response protein, Dps family - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: B75628
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MID:20036896; PMID:10567266
 A:Accession: B75628
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-241 <WHI>
 A:Cross-references: GB:AE001826; MID:96460827; PIDN:AAF12541.1; PID:96460837; TIGR:DR
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRB0092
 A:Map position: megaplasmid
 A:Genome: plasmid
 A:Note: plasmid MPI
 Query Match 0.9%; Score 9; DB 2; Length 241;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 419 VVVVSSILL 427
 DB 8 VVVVSSILL 16
 RESULT 5
 AB83436
 hypothetical protein PA1680 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: AB83436
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: AB2950; MID:20437337; PMID:10984043
 A:Accession: AB83436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-327 <STO>
 A:Cross-references: GB:AE004595; GB:AE004091; MID:99947644; PIDN:AAG05069.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA1680
 Query Match 0.9%; Score 9; DB 2; Length 327;
 Best Local Similarity 100.0%; Pred. No. 2.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 340 LRLRAEALA 348
 DB 81 LRLRAEALA 89
 RESULT 6
 T24380
 hypothetical protein T03D8.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T24380
 R:Montmore, B.

submitted to the EMBL Data Library, March 1997
 A:Reference number: 219884
 A:Accession: T24380
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-157 <WIL>
 A:Cross-references: EMBL:292838; PIDN:CAB07406.1; GSPDB:GN00023; CESP:T03D8.2
 A:Experimental source: clone T03D8
 C:Genetics:
 A:Gene: CESP:T03D8.2
 A:Map position: 5
 A:Introns: 49/3; 83/1; 125/3

Query Match 0.8%; Score 8; DB 2; Length 157;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 501 TLETHQSS 508
 DB 27 TLETHQSS 34
 |||||

RESULT 7
 E90194
 hypothetical protein psta [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 A:Accession: E90194
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
 arett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 A:Submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: E90194
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-272 <KUR>
 A:Cross-references: GB:AE006641; NID:g13813648; PIDN:AAK40812.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: psta

Query Match 0.8%; Score 8; DB 2; Length 272;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1015 QISEALRI 1022
 DB 152 QISEALRI 159
 |||||

RESULT 8
 G95117
 HemK protein [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 A:Accession: G95117
 R:Petelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held-
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Kadane, D.; Holtzapfel,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: G95117
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-279 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75136.1; PID:g14972494; GSPDB:GN00164; TIGR:SP4
 C:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP1021

Query Match 0.8%; Score 8; DB 2; Length 279;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 LVELLIAE 581
 DB 99 LVELLIAE 106
 |||||

RESULT 9
 E88103
 protein W10G11.5 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 A:Accession: E88103
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio-
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
 A:Accession: E88103
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-345 <STO>
 A:Cross-references: GB:chr_II; PIDN:AB95070.1; PID:g2746928; GSPDB:GN00020; CESP:W10G
 A:Gene: W10G11.5
 A:Map position: 2

Query Match 0.8%; Score 8; DB 2; Length 345;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 683 TLLADTSD 690
 DB 50 TLLADTSD 57
 |||||

RESULT 10
 G89801
 hypothetical protein SA0346 [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 A:Accession: G89801
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O-
 ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: G89801
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-386 <KUR>
 A:Cross-references: GB:BA000018; PID:g13700272; PIDN:BA841570.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: SA0346
 C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 0.8%; Score 8; DB 2; Length 386;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 LPLVSVSL 877
 DB 312 LPLVSVSL 319
 |||||

RESULT 11
 A29393
 hemocyanin beta-c chain d - Roman snail

C:Species: *Helix pomatia* (Roman snail)
 C:Date: 13-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 11-Jan-2000
 C:Accession: A29393
 R:Prexell, R.; Slegmund, S.; Schneider, H.J.; Linzen, B.; Gielens, C.; Preaux, G.; Lontje
 Biol. Chem. Hoppe-Seyler 368, 617-635, 1987
 A:Title: Complete amino-acid sequence of a functional unit from a molluscan hemocyanin
 A:Reference number: A29393; MUID:87299002; PMID:3620107
 A:Accession: A29393
 A:Molecule type: protein
 A:Residues: 1-410 <DRE>
 C:Superfamily: hemocyanin

Query Match 0.88; Score 8; DB 2; Length 410;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 523 SREORFDP 530
 |||||
 DB 115 SREORFDP 122

RESULT 12

A70947
 hypothetical protein RV3163c - *Mycobacterium tuberculosis* (strain H37RV)

C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: A70947
 R:Coyle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 Rajandream, M.A.; Davies, R.; Devlin, R.; Fellner, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
 Nature 393, 537-544, 1998
 A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: A70947
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-423 <COL>
 A:Cross-references: GB:AL021646; GB:AL123456; NID:93242278; PIDN:CAAL6628.1; PID:9282757
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3163c

Query Match 0.88; Score 8; DB 2; Length 423;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 975 VVOTALRN 982
 |||||
 DB 258 VVOTALRN 265

RESULT 13

AB0046
 probable N-acetylmuramoyl-L-alanine amidase-family protein YPO0370 [imported] - *Yersinia*

C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002
 C:Accession: AB0046
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchall, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 et al.; M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AB0046
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-593 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CA089229.1; PID:91597468; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO0370
 C:Superfamily: *Escherichia coli* hypothetical protein b2817

Query Match 0.88; Score 8; DB 2; Length 593;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 417 KPVVWSS 424
 |||||
 DB 168 KPVVWSS 175

RESULT 14

F75553
 hypothetical protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
 C:Accession: F75553
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: F75553
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-640 <MID>
 A:Cross-references: GB:AE001878; GB:AE000513; NID:96457820; PIDN:AAF09751.1; PID:9645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR0160
 A:Map position: 1
 C:Superfamily: *Deinococcus radiodurans* hypothetical protein DR0160

Query Match 0.88; Score 8; DB 2; Length 640;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 840 LLAGLSP 847
 |||||
 DB 372 LLAGLSP 379

RESULT 15

C87504
 hypothetical protein CC2058 [imported] - *Caulobacter crescentus*

C:Species: *Caulobacter crescentus*
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: C87504
 R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; Debey, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: C87504
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-958 <STO>
 A:Cross-references: GB:AE005673; NID:913423537; PIDN:AAK24031.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC2058

Query Match 0.88; Score 8; DB 2; Length 958;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 328 SAGRRSD 335
 |||||
 DB 665 SAGRRSD 672

Search completed: May 7, 2003, 19:46:23
 Job time: 33 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:42:23 ; Search time 15 Seconds
(without alignments)
2845.276 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 1029
Sequence: 1 MHILVHAMVILLTLGPPRA.....MDPSAQISEALRIHMEAVM 1029

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	9	188	1 R1MM_DEIRA	O9rsu1 delinococcus
2	8	335	1 LEP4_MYXXA	O30387 myxococcus
3	8	338	1 CYF_PHOIA	P93522 heliobacterium
4	8	410	1 HCYB_HEIPO	P12031 helix pomati
5	7	128	1 GIBO_MYCTU	O53197 mycobacteri
6	7	163	1 BPB1_AMPPE	O17077 anthracis p
7	7	172	1 ALI2_BOVIN	O28133 bos taurus
8	7	197	1 YRFP_ECOLI	P75547 escherichia
9	7	207	1 RECR_RALSO	O8y050 ralsionia s
10	7	208	1 SPC3_STRPU	P16537 strongyloce
11	7	209	1 ENMD_ECO57	O8xbw8 escherichia
12	7	209	1 ENMD_ECOLI	P19925 escherichia
13	7	209	1 ENMD_SHITL	O54153 shigella fl
14	7	211	1 THIE_BACHD	O9kcy8 bacillus ha
15	7	233	1 Y33L_SYNT3	P72583 synchocyst
16	7	244	1 HIS4_BACHD	O9k625 bacillus ha
17	7	245	1 OCCM_RHIME	P72296 rhizobium m
18	7	258	1 NORC_NEIMA	O91vo0 neisseria m
19	7	262	1 ATP6_ALLAR	P50363 allomyces a
20	7	262	1 ATP6_ALUMA	P50364 allomyces m
21	7	264	1 COMT_RAT	P22734 rattus norv
22	7	265	1 COMT_MOUSE	O88587 mus musculu
23	7	270	1 RPA2_MOUSE	O62193 mus musculu
24	7	285	1 YX47_STRCO	O9x8h0 streptomyce
25	7	287	1 DCMH_HYDPS	P19914 hydrogroph
26	7	289	1 FRAH_ANASP	P46017 anabaena sp
27	7	289	1 KPPR_ECOLI	P37307 escherichia
28	7	290	1 VINT_BPL2	P42540 bacterioph
29	7	290	1 BCHL_RHORI	O918j5 rhodospiril
30	7	293	1 VNP7_CAEBL	P34560 caenorhabdi
31	7	294	1 SC17_PICPA	O9p4d0 plinia past
32	7	299	1 YRBE_ECOLI	P75826 escherichia
33	7	300	1 HAP2_KLUDA	P53768 kluyveromyc

ALIGNMENTS

```

RESULT 1
R1MM_DEIRA          STANDARD:      PRT: 188 AA.
AC O9RSU1:
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 16S rRNA processing protein r1mm.
GN R1MM OR DR2010.
OS Delinococcus radiodurans.
OC Bacteria; Thermus/Delinococcus group; Delinococci; Delinococcales;
OC Delinococcaceae; Delinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.U., Hatt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Delinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -!- FUNCTION: ESSENTIAL FOR EFFICIENT PROCESSING OF 16S RNA. PROBABLY
CC PART OF THE 30S SUBUNIT PRIOR TO OR DURING THE FINAL STEP IN THE
CC PROCESSING OF 16S PRE-30S RIBOSOMAL SUBUNITS. IT COULD BE SOME
CC ACCESSORY PROTEIN NEEDED FOR EFFICIENT ASSEMBLY OF THE 30S
CC SUBUNIT. R1MM IS NEEDED IN A STEP PRIOR TO RBEA DURING THE
CC MATURATION OF 16S RNA. HAS AFFINITY FOR FREE RIBOSOMAL 30S
CC SUBUNITS BUT NOT FOR 70S RIBOSOMES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -!- SIMILARITY: BELONGS TO THE R1MM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isp.slb.ch/announce/
CC or send an email to license@slb.slb.ch).
CC -----
CC EMBL: AE002038; AAF11557.1; -
CC TIGR: DR2010; -
CC DR InterPro: IPR002676; R1MM.
CC DR Pfam: PF01782; R1MM; 1.
CC RNA processing; Complete proteome.
CC SO SEQUENCE 188 AA; 19686 MW; 495828D781CC99FD CRC64;
Query Match 0.9%; Score 9; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 0.65; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

```

OY 432 PLAGKPGA 440
 DB 155 PLAGKPGA 163

RESULT 2

LEP4_MYXA
 ID LEP4_MYXA STANDARD: PRT: 335 AA.
 AC 030387;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Type 4 prepillin-like proteins leader peptide processing enzyme
 DE (includes: Leader peptidase (EC 3.4.99.-) (Pepillin peptidase); N-
 methyltransferase (EC 2.1.1.-)).
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 OC Myxococcales; Cytophactinaceae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=DK1622;
 RA Wu S.S., Cheng Y.L., Kaiser D.;
 RT "The pilH gene encodes an ABC transporter required for type IV pilus
 RT biogenesis in Myxococcus xanthus."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CLEAVES TYPE-4 FIMBRIAL LEADER SEQUENCE AND METHYLATES
 CC THE N-TERMINAL (GENERALLY PHE) RESIDUE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A24.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF003632; AAC36157.1;
 DR MEROPS: A24.001;
 DR InterPro: IPR000045; Peptidase_C20.
 DR Pfam: PF01478; Peptidase_C20.1.
 DR PRINTS: PR00864; PREPLINPTASE.
 DR MultiFunctional enzyme: Hydrolase; Protease; Transferase;
 KW Methyltransferase; Transmembrane.
 FT TRANSMEM 13 33 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT TRANSMEM 258 278 POTENTIAL.
 FT TRANSMEM 299 319 POTENTIAL.
 SQ SEQUENCE 335 AA; 36767 MW; 27DF58A5FD2819FA CRC64;
 Query Match 0.8%; Score 8; DB 1; Length 335;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 857 LDEGEER 864
 DB 283 LDEGEER 290

RESULT 3

CYF_PHOLA
 ID CYF_PHOLA STANDARD: PRT: 338 AA.
 AC P95522;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Apocytochrome F precursor.
 GN PETA.

OS Phormidium lamosum.
 OC Bacteria; Cyanobacteria; Oscillatoriales; Phormidium.
 OX NCBI_TaxID=32059;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97009001; PubMed=8856106;
 RA Wagner M.J., Packer J.C.L., Howe C.J., Bendall D.S.;
 RT "Some characteristics of cytochrome f in the cyanobacterium Phormidium
 RT lamosum: its sequence and charge properties in the reaction with
 RT plastocyanin."
 RL Biochim. Biophys. Acta 1276:246-252(1996).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 46-295.
 RX MEDLINE=99352179; PubMed=10423236;
 RA Carrell C.J., Schlarb B.G., Bendall D.S., Howe C.J., Cramer W.A.,
 RA Smith J.L.;
 RT "Structure of the soluble domain of cytochrome f from the
 RT cyanobacterium Phormidium lamosum."
 RL Biochemistry 38:9590-9599(1999).
 CC -1- FUNCTION: TRANSLOCATES PROTONS ACROSS THE THYLAKOID MEMBRANE AND
 CC RECEIVES ELECTRONS FROM THE RIESKE IRON-SULFUR PROTEIN AND PASSES
 CC THEM TO PLASTOCYANIN. THIS FUNCTION IS VERY SIMILAR TO THAT OF
 CC MITOCHONDRIAL CYTOCHROME C1.
 CC -1- SUBUNIT: MEMBER OF THE CYTOCHROME B6/F COMPLEX INCLUDING
 CC CYTOCHROME B6, CYTOCHROME F AND PROBABLY AN IRON SULFUR PROTEIN.
 CC -1- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Y09612; CAI70824.1;
 DR PDB: 1CI3; 3I-MAY-00.
 DR InterPro: IPR002325; Apocyt_F.
 DR Pfam: PF01333; Apocytochrome_F.1.
 DR PRINTS: PS00180; CYTOCHROME_C.1.
 DR PROSITE: PS00180; CYTOCHROME_C.1.
 KW Electron transport; Heme; Photosynthesis; Photosystem I;
 KW Photosystem II; Signal; 3D-structure.
 FT SIGNAL 1 45
 FT CHAIN 46 338 APOCYTOCHROME F.
 FT METAL 46 46 IRON (HEME AXIAL LIGAND).
 FT BINDING 66 66 HEME (COVALENT).
 FT BINDING 69 69 HEME (HEME AXIAL LIGAND).
 FT METAL 70 70 IRON (HEME AXIAL LIGAND).
 SQ SEQUENCE 338 AA; 36069 MW; BD3D52036181FB50 CRC64;
 Query Match 0.8%; Score 8; DB 1; Length 338;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 9 MWILLTG 16
 DB 23 MWILLTG 30

RESULT 4

HCBY_HELPO
 ID HCBY_HELPO STANDARD: PRT: 410 AA.
 AC P12031;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hemocyanin, beta-C chain unit D (Fragment).
 OS Helix pomatia (Roman snail) (Edible snail).
 GN Eukaryota; Metazoa; Mollusca; Gastropoda; Stylomatophora;

OC Helicaceae; Helicidae; Helix.
 OX NCBI_TaxID=5336;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=87299002; PubMed=3620107;
 RA Drexel R., Siegmund S., Schneider H.J., Linzen B., Gjelens C.,
 RA Preaux G., Lontle R., Kellermann J., Lottspeich F.;
 RT "Complete amino-acid sequence of a functional unit from a molluscan
 RT haemocyanin (Helix pomatia).";
 RL Biol. Chem. Hoppe-Seyler 368:617-633(1987).
 RN [2]
 RP PARTIAL SEQUENCE, THIOETHER BOND, REVISIONS, AND MASS SPECTROMETRY.
 RX MEDLINE=98000289; PubMed=9342242;
 RA Gjelens C., de Geest N., Xin X.-Q., Devreese B., van Beeumen J.,
 RA Preaux G.;
 RT "Evidence for a cysteine-histidine thioether bridge in functional
 RT units of molluscan haemocyanins and location of the disulfide bridges
 RT in functional units d and g of the beta-c-haemocyanin of Helix
 RT pomatia.";
 RL Eur. J. Biochem. 248:879-888(1997).
 CC -1- FUNCTION: HEMOCYANINS ARE COPPER-CONTAINING OXYGEN CARRIERS
 CC OCCURRING FREELY DISSOLVED IN THE HEMOLYMPH OF MANY MOLLUSKS AND
 CC ARTHROPODS.
 CC -1- COFACTOR: CONTAINS TWO COPPER ATOMS PER FUNCTIONAL UNIT.
 CC -1- SUBUNIT: DECAMERS OF LARGE IDENTICAL SUBUNITS (450 kDa), EACH
 CC CONTAINING 8 GLOBULAR OXYGEN-BINDING FUNCTIONAL UNITS.
 CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY. HEMOCYANIN
 CC SUBFAMILY.
 CC PTR: A29393; A29393.
 DR InterPro: IPR002227; Tyrosinase.
 DR Pfam: PF00264; Tyrosinase. 1.
 DR PRINTS: PR00092; TYROSINASE.
 DR PROSITE: PS00497; TYROSINASE_1; 1.
 DR PROSITE: PS00498; TYROSINASE_2; 1.
 KW Oxygen transport; Transport; Copper; Glycoprotein;
 KW Hemolymph; Repeat.
 FT METAL 1 1
 FT METAL 44 44
 FT METAL 55 55
 FT METAL 71 71
 FT METAL 175 175
 FT METAL 179 179
 FT METAL 206 206
 FT DISULFID 50 59
 FT DISULFID 165 232
 FT DISULFID 321 332
 FT THIOETH 60 62
 FT CARBOHYD 253 253
 FT CONFLICT 60 60
 FT CONFLICT 62 62
 FT CONFLICT 165 165
 FT NON_TER 410 410
 SQ SEQUENCE 410 AA; 47019 MW; E872743ED1AE66BF CRC64;
 Query Match 0.8%; Score 8; DB 1; Length 410;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 523 SREORFDP 530
 DB 115 SREORFDP 122
 ID GIBO_MYCTU STANDARD; PRT; 128 AA.
 AC 053197;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hemoglobin-like protein HBO.
 GN GIBO OR RV2470 OR MT2546 OR MTV008.26.
 OS Mycobacterium tuberculosis, and

OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773, 1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Baevkov K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gents S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellon M., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Hatt D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.bovis; STRAIN=BCG;
 RA Pathania R., Ramande P., Dikshit K.L.;
 RT "Cloning, expression and characterization of an oxygen-binding protein
 RT encoded by gibo locus of Mycobacterium tuberculosis and Mycobacterium
 RT bovis.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE PROTOZOAN/CYANOBACTERIAL GLOBIN
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial/
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AL021246; AAA16047.1; .
 CC EMBL: AE007091; AAK46846.1; ALT_INIT.
 CC EMBL: AF213450; AAF36597.1; .
 CC TIGR: MT2546; .
 DR TuberculList; RV2470; .
 DR InterPro: IPR001486; Bac_globin.
 DR ProDom: PD004840; Bac_globin; 1.
 DR PROSITE: PS01213; GLOBIN_FAM_2; 1.
 KW Heme; Oxygen transport; Transport; Complete proteome.
 FT METAL 75 75
 FT METAL 75 75
 SQ SEQUENCE 128 AA; 14950 MW; CEF4CD23907B0CCA CRC64;
 Query Match 0.7%; Score 7; DB 1; Length 128;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 187 PEDDIAG 193
 DB 37 PEDDIAG 43
 ID PBP1_ANTPE STANDARD; PRT; 163 AA.
 AC 053197;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hemoglobin-like protein HBO.
 GN GIBO OR RV2470 OR MT2546 OR MTV008.26.
 OS Mycobacterium tuberculosis, and

AC 017077;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pheromone-binding protein 1 precursor (PBP 1) (APR-1).
 OS Antherea pernyi (Chinese oak silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 OC Diptysia; Bombycoidea; Saturniidae; Saturniinae; Saturniini;
 OC Antherea.
 RX NCBI_TaxID=7119;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Antenna;
 RX MEDLINE=91268328; PubMed=2097299;
 RA Raming K., Krieger J., Breer H.;
 RT "Primary structure of a pheromone-binding protein from Antherea
 pernyi: homologues with other ligand-carrying proteins.";
 RL J. Comp. Physiol. B 160:503-509(1990).
 CC -1- FUNCTION: THIS MAJOR SOLUBLE PROTEIN IN OLFACTORY SENSILLA OF MALE
 MOTHS MIGHT SERVE TO SOLUBILIZE THE EXTREMELY HYDROPHOBIC
 PHEROMONE MOLECULES AND TO TRANSPORT PHEROMONE THROUGH THE AQUEOUS
 LYMPH TO RECEPTORS LOCATED ON OLFACTORY CILIA.
 CC -1- TISSUE SPECIFICITY: ANTENNA.
 CC -1- SIMILARITY: BELONGS TO THE PBP/GOBP FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X96773; CAA65576.1; -
 DR HSSP: P34174; IDOE.
 DR InterPro: IPR000746; PBP_GOBP_family.
 DR Pfam: PF01395; PBP_GOBP, 1.
 DR PRINTS: PR00484; PBPGOBP.
 KW Pheromone-binding; Pheromone response; Transport; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 163 163 PHEROMONE-BINDING PROTEIN 1.
 SO SEQUENCE 163 AA; 18086 MW; 3CFE4FE3E369C56 CRC64;
 QY 698 GADASMA 704
 DB 102 GADASMA 108
 Query Match 0.7%; Score 7; DB 1; Length 163;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RT "Complementary DNA cloning of the predominant allergen of bovine
 dander: a new member in the lipocalin family.";
 RL J. Allergy Clin. Immunol. 97:1297-1303(1996).
 RN [2]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=99068569; PubMed=9853680;
 RA Santa H., Saarela J.T., Laatikainen R., Rautalahti J., Virtanen T.,
 RA Rytönen M., Maentylä R.;
 RT "A bovine dander allergen, comparative modeling, and similarities and
 differences in folding with related proteins.";
 RL J. Protein Chem. 17:657-662(1998).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ÅNGSTRÖMS).
 RX MEDLINE=99107890; PubMed=9891000;
 RA Rouvainen J., Rautalahti J., Virtanen T., Zeller T., Kauppinen J.,
 RA Taivainen A., Maentylä R.;
 RT "Probing the molecular basis of allergy, three-dimensional structure
 of the bovine lipocalin allergen bos d 2.";
 RL J. Biol. Chem. 274:2337-2343(1999).
 RN [4]
 RP CHARACTERIZATION, AND CRYSTALLIZATION.
 RX MEDLINE=98321199; PubMed=9647765;
 RA Rautalahti J., Auriole S., Rouvainen J., Kauppinen J., Zeller T.,
 RA Novikov D., Virtanen T., Maentylä R.A.;
 RT "Molecular and crystal properties of bos d 2, an allergenic protein of
 the lipocalin family.";
 RL Biochem. Biophys. Res. Commun. 247:746-750(1998).
 RN [5]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=98184363; PubMed=9525451;
 RA Rautalahti J., Rytönen M., Syrjänen K., Penttinen J., Zeller T.,
 RA Virtanen T., Maentylä R.;
 RT "Tissue localization of bovine dander allergen Bos d 2.";
 RL J. Allergy Clin. Immunol. 101:349-353(1998).
 CC -1- FUNCTION: PROBABLE PHEROMONE CARRIER.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN SKIN. PRODUCED IN SWEAT
 GLANDS AND TRANSPORTED TO THE SKIN SURFACE.
 CC -1- MISCELLANEOUS: PREDOMINANT ALLERGEN OF BOVINE DANDER.
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: I42867; AAB08720.1; -
 DR PDB: 1BUT; 1I-MAY-99.
 DR InterPro: IPR000566; Lipocalin_cytfabp.
 DR Pfam: PF00061; Lipocalin, 1.
 DR PROSITE: PS00213; LIPOCALIN; FALSE-NEG.
 KW Transport; Lipocalin; Signal; Allergen; 3D-structure.
 FT SIGNAL 1 16
 FT CHAIN 17 172 ALLERGEN BOS D 2.
 FT MOD_RES 17 17 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 60 64
 FT DISULFID 79 170
 SO SEQUENCE 172 AA; 19560 MW; C50F66F1B3D98156 CRC64;
 QY 311 VEGGPLR 317
 DB 46 VEGGPLR 52
 Query Match 0.7%; Score 7; DB 1; Length 172;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8
 YFPF_ECOLI

```

ID YFPF_ECOLI STANDARD: PRT: 197 AA.
AC P76547;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yfpf.
GN YFPF OR B2447.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia
CC NCBI_TaxID=562;
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000331; AAC75500.1; -.
DR EcoGene; EG14177; yfpf.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 197 AA; 22907 MW; 6E1BE6A72D49B519 CRC64;

Query Match 0.7%; Score 7; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 821 ADPLHDL 827
DB 15 ADPLHDL 21

RESULT 9
RECR_RALSO STANDARD: PRT: 207 AA.
AC Q87050;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Recombination protein recr.
GN RECR OR RSC1194 OR RS05726.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
CC Ralstonia.
CC NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billaut A., Broctier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Gunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Stigter P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
CC -1- FUNCTION: May play a role in DNA repair. It seems to be involved
CC in an recBC-independent recombinational process of DNA repair. It
CC may act with recF and recO (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RECR FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL646063; CAD14896.1; -.
DR InterPro; IPR002936; DNAPrim_toprim.
DR InterPro; IPR000093; RecR.
DR Pfam; PF01751; Toprim; 1.
DR Pfam; PF02132; RecR; 1.
DR SMART; SM00493; TOPRIM; 1.
DR TIGRfams; TIGR00615; recr; 1.
DR PROSITE; PS01300; RECR; FALSE_NEG.
KW DNA repair; DNA recombination; Zinc-finger; Complete proteome.
FT ZN_FING 62
FT C4_TYPE (POTENTIAL).
SQ SEQUENCE 207 AA; 22510 MW; 1428A8F817D03071 CRC64;

Query Match 0.7%; Score 7; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 342 RLAEALA 348
DB 47 RLAEALA 53

RESULT 10
SPEC3_STRPU STANDARD: PRT: 208 AA.
AC P16537;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE SPEC3 protein.
GN SPEC3.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
CC Strongylocentrotus.
CC NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88112804; PubMed=2828169;
RA Eldon E.D., Angerer L.M., Angerer R.C., Klein W.H.;
RT "Spec3: embryonic expression of a sea urchin gene whose product is
RT involved in ectodermal ciliogenesis."
RL Genes Dev. 1:1280-1292(1987).
RN [2]
RP SUBCELLULAR LOCATION.
RX MEDLINE=90169510; PubMed=2407617;
RA Eldon E.D., Montpelit I.C., Nguyen T., Decker G., Valdivia M.C.,
RA Klein W.H., Brandhorst B.P.;
RT "Localization of the sea urchin Spec3 protein to cilia and Golgi
RT complexes of embryonic ectoderm cells."
RL Genes Dev. 4:111-122(1990).
CC -1- SUBCELLULAR LOCATION: IN CILIA AND GOLGI COMPLEXES.
DR PIR; A43696; A43696.
KW Transmembrane.
FT TRANSMEM 107 127
FT TRANSMEM 163 183
FT CAROXYD 71 71
SQ SEQUENCE 208 AA; 21719 MW; 4BE0638F1DE65C2E CRC64;

Query Match 0.7%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 438 PCADGGS 444
DB 11111111

```

Db 138 PGADGS 144

RESULT 11

ENTD_ECOS7 STANDARD: PRT; 209 AA.

ID ENTD_ECOS7

AC Q8XBH8;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 4'-phosphopantetheinyl transferase end (EC 2.7.8.-) (Enterobactin synthetase component D) (Enterobactin synthase D).

GN ENTD OR Z0723 OR ECS0622.

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=83334;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; PubMed=11206551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai J., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;

RA "Genome sequence of enterohemorrhagic Escherichia coli O157:H7"; Nature 409:529-533(2001).

RL [2]

RN SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / K1MD 0509952;

RX PubMed=11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kuhara S., Shiba T., Hattori M., Shingawa H.;

RA "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12."; DNA Res. 8:11-22(2001).

RL -1- FUNCTION: Catalyzes the transfer of the 4'-phosphopantetheine moiety from coenzyme A to apo-domains of both entB (an ATP domain) and entF (a PCP domain). Plays an essential role in the assembly of the enterobactin (By similarity).

CC -1- CATALYTIC ACTIVITY: CoA + apo-entB/F -> adenosine 3',5'-bisphosphate + holo-entB/F.

CC -1- COFACTOR: Magnesium (By similarity).

CC -1- PATHWAY: Enterobactin biosynthesis.

CC -1- SUBUNIT: EntB, entD, entE, and entF form a multienzyme complex called enterobactin synthetase (By similarity).

CC -1- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY. ENTD FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE005238; AAG54917.1; -

DR EMBL: AF002352; BAB34045.1; -

DR InterPro: IPR002582; ACPS.

DR InterPro: IPR003542; Enbac-synthetD.

DR Pfam: PF01648; ACPS; 1.

DR PRINTS: PR01399; ENTSTHTASD.

KW Transferrase; Enterobactin biosynthesis; Transport; Iron transport; Magnesium; Complete proteome.

KW METAL

FT METAL 110 110 MAGNESIUM (BY SIMILARITY).

FT METAL 112 112 MAGNESIUM (BY SIMILARITY).

FT METAL 155 155 MAGNESIUM (BY SIMILARITY).

SEQUENCE 209 AA; 23477 MW; 236960055E5765AE CRC64;

Query Match 0.7%; Score 7; DB 1; Length 209;

Best Local Similarity 100.0%; Pred. No. 74;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 941 LAFSLAL 947

Db 141 LAFSLAL 147

RESULT 12

ENTD_ECOLI STANDARD: PRT; 209 AA.

ID ENTD_ECOLI

AC P19925; P77092;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 4'-phosphopantetheinyl transferase end (EC 2.7.8.-) (Enterobactin synthetase component D) (Enterobactin synthase D).

GN ENTD OR B0583.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K12;

RX MEDLINE=90132534; PubMed=2533240;

RA Codere P.E., Earhart C.F.;

RA "The entD gene of the Escherichia coli K12 enterobactin gene cluster."; J. Gen. Microbiol. 135:3043-3055(1989).

RL [2]

RN ERRATUM.

RA Codere P.E., Earhart C.F.;

RA J. Gen. Microbiol. 136:1667-1667(1990).

RL [3]

RN SEQUENCE FROM N.A.

RP MEDLINE=89313305; PubMed=2526281;

RA Armstrong S.K., Pettis G.S., Forrester L.J., McIntosh M.A.;

RA "The Escherichia coli enterobactin biosynthesis gene, entD; nucleotide sequence and membrane localization of its protein product."; Mol. Microbiol. 3:757-766(1989).

RL [4]

RN SEQUENCE FROM N.A.

RP STRAIN-K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).

RL [5]

RN SEQUENCE FROM N.A.

RC STRAIN-K12;

RX MEDLINE=97061202; PubMed=8905223;

RA Oshino K., Iida T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizochichi K., Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N., Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;

RA "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map."; DNA Res. 3:137-155(1996).

RL [6]

RN SEQUENCE FROM N.A.

RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M., Federspiel N., Hyman R., Kaiman S., Komp C., Kudl O., Lew H., Lin D., Nannath A., Oefner P., Roberts D., Schramm S., Davis R.W.;

RA Submitted (JAN-1997) to the EMBL/GenBank/DBD databases.

```

RN [7]
RP FUNCTION.
RX PubMed-8999709;
RA Lambiot R.H., Gehring A.M., Flugel R.S., Zuber P., Lacelle M.,
RA Marahel M.A., Reid R., Khosla C., Walsh C.T.;
RT "A new enzyme superfamily - the phosphopantetheinyl transferases.";
RL Chem. Biol. 3:923-936(1996).
RN [8]
RP FUNCTION.
RX STRAIN-BL21-DE3;
RX PubMed-9214294;
RA Gehring A.M., Bradley K.A., Walsh C.T.;
RT "Enterobactin biosynthesis in Escherichia coli: isochorismate lyase
RT (Entb) is a bifunctional enzyme that is phosphopantetheinylated by
RT Entb and then acetylated by Ente using ATP and 2,3-dihydroxybenzoate.";
RL Biochemistry 36:8495-8503(1997).
CC -1- FUNCTION: Catalyzes the transfer of the 4'-phosphopantetheine
CC moiety from coenzyme A to apo-domains of both entb (an ArCP
CC domain) and entf (a PCP domain). Plays an essential role in the
CC assembly of the enterobactin.
CC -1- CATALYTIC ACTIVITY: CoA + apo-entb/F = adenosine 3',5'-
CC bisphosphate + holo-entb/F.
CC -1- COFACTOR: Magnesium.
CC -1- PATHWAY: Enterobactin biosynthesis.
CC -1- SUBUNIT: Entb, entd, ente, and entf form a multienzyme complex
CC called enterobactin synthetase.
CC -1- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY.
CC ENTD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17426; CAB57861.1; -
DR EMBL; AE000163; AAC73684.1; -
DR EMBL; D90700; BAA35224.1; ALT_INIT.
DR EMBL; U82598; AAB40782.1; ALT_INIT.
DR PIR; S06981; S06981.
DR Ecogen; EG10262; entd.
DR InterPro: IPR002582; ACPS.
DR InterPro: IPR003542; Embac_synthetd.
DR Pfam: PF01648; ACPS.1.
DR PRINTS: PR01399; ENTSNTHASED.
KW Transferase; Enterobactin biosynthesis; Transport; Iron transport;
KW Magnesium; Complete proteome.
FT METAL 110 110 MAGNESIUM (BY SIMILARITY).
FT METAL 112 112 MAGNESIUM (BY SIMILARITY).
FT METAL 155 155 MAGNESIUM (BY SIMILARITY).
FT METAL 77 78 EL-> DV (IN REF. 3)
FT CONFLICT 77 78
SQ SEQUENCE 209 AA; 23604 MW; B98403563D9DA7D CRC64;

Query Match 0.7%; Score 7; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 941 LAFSIAL 947
DB 141 LAFSIAL 147

```

```

GN ENTD.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RA Johansen K.A.;
RT "entd enterobactin biosynthesis genes of enteric bacteria.";
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Catalyzes the transfer of the 4'-phosphopantetheine
CC moiety from coenzyme A to apo-domains of both entb (an ArCP
CC domain) and entf (a PCP domain). Plays an essential role in the
CC assembly of the enterobactin (BY similarity).
CC -1- CATALYTIC ACTIVITY: CoA + apo-entb/F = adenosine 3',5'-
CC bisphosphate + holo-entb/F.
CC -1- COFACTOR: Magnesium (BY similarity).
CC -1- PATHWAY: Enterobactin biosynthesis.
CC -1- SUBUNIT: Entb, entd, ente, and entf form a multienzyme
CC complex called enterobactin synthetase (BY similarity).
CC -1- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY.
CC ENTD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U52684; AAA97936.1; -
DR InterPro: IPR002582; ACPS.
DR InterPro: IPR003542; Embac_synthetd.
DR Pfam: PF01648; ACPS.1.
DR PRINTS: PR01399; ENTSNTHASED.
KW Transferase; Enterobactin biosynthesis; Transport; Iron transport;
KW Magnesium.
FT METAL 110 110 MAGNESIUM (BY SIMILARITY).
FT METAL 112 112 MAGNESIUM (BY SIMILARITY).
FT METAL 155 155 MAGNESIUM (BY SIMILARITY).
FT METAL 77 78
SQ SEQUENCE 209 AA; 23487 MW; 72F2B8C05C19C3E CRC64;

Query Match 0.7%; Score 7; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 941 LAFSIAL 947
DB 141 LAFSIAL 147

RESULT 14
THIE_BACHD
ID THIE_BACHD STANDARD; PRT; 211 AA.
AC Q9KCY8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thiamine-phosphate pyrophosphorylase (EC 2.5.1.3) (TMP
DE pyrophosphorylase) (TMP-Pase) (Thiamine-phosphate synthase).
GN THIE OR BH1431.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-C-125 / JCM 9153;
RA MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kihara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus

```

RT halodurans and genomic sequence comparison with *Bacillus subtilis*:"
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- FUNCTION: Condenses 4-methyl-5-(beta-hydroxyethyl)-thiazole
 CC monophosphate (THZ-P) and 4-amino-5-hydroxymethyl pyrimidine
 CC pyrophosphate (HMP-PP) to form thiamine monophosphate (TMP) (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: 2-methyl-4-amino-5-hydroxymethylpyrimidine
 CC diphosphate + 4-4-methyl-5-(2-phosphonooxyethyl)-thiazole -
 CC diphosphate + thiamine monophosphate.
 CC -1- PATHWAY: Thiamine biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE TMP-PPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP001512; BAB05150.1; -
 CC DR HSSP: P39594; 2TPS.
 CC DR InterPro: IPR003733; TMP_synthase.
 CC DR pfam: PF02581; TMP-TEN1.1
 CC DR TIGRfam: TIGR00693; thtP.1
 CC KW Thiamine biosynthesis; Transferase; Complete proteome.
 CC SEQUENCE 211 AA; 23150 MW; 86AAB50608D2D6E CRC64;
 SQ
 Query Match 0.7%; Score 7; DB 1; Length 211;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 243 RVLQALA 249
 DB 52 RVLQALA 58
 RESULT 15
 Y53L_SYNV3
 ID Y53L_SYNV3 STANDARD; PRT; 233 AA.
 AC P72583;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ycf53-like protein.
 GN Ycf53.
 GN SL0558.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID-1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97061201; PubMed-8905231;
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugliara M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nanno K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.,
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1- SIMILARITY: BELONGS TO THE YCF53 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D90899; BAA16582.1; -
 CC DR Hypothetical protein; Complete proteome.
 KW

SQ SEQUENCE 233 AA; 26465 MW; DE36B048AC83757C CRC64;
 Query Match 0.7%; Score 7; DB 1; Length 233;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 968 LGSQDFE 974
 DB 100 LGSQDFE 106
 Search completed: May 7, 2003, 19:44:54
 Job time : 19 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:43:49 ; Search time 19 Seconds
(without alignments)
1593.484 Million cell updates/sec

Title: US-09-929-769-7

Perfect score: 1029

Sequence: 1 MHILVHAMVILLTGPPRA.....MDPSAQISEALRIHMEAVM 1029

Scoring table: OLIGO

Gapop 60.0 , Gapect 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRMUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfill.esl.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	0.8	224	1	US-08-608-241-4
2	8	0.8	224	2	US-08-922-182-4
3	8	0.8	224	2	US-08-919-953-4
4	8	0.8	226	4	US-09-192-983-4
5	8	0.8	390	4	US-09-308-003-12
6	8	0.8	396	4	US-09-134-001C-4580
7	7	0.7	27	2	US-08-660-789-9
8	7	0.7	27	4	US-09-074-114-9
9	7	0.7	30	1	US-08-127-351-8
10	7	0.7	30	1	US-08-480-367B-8
11	7	0.7	30	1	US-08-487-221A-8
12	7	0.7	30	1	US-08-480-370-8
13	7	0.7	33	1	US-08-127-351-9
14	7	0.7	33	1	US-08-480-367B-9
15	7	0.7	33	1	US-08-487-221A-9
16	7	0.7	33	1	US-08-480-370-9
17	7	0.7	35	4	US-09-460-145-9
18	7	0.7	291	4	US-09-286-691-28
19	7	0.7	291	4	US-09-687-147-28
20	7	0.7	317	6	5340934-11
21	7	0.7	331	4	US-09-061-702-4
22	7	0.7	331	4	US-09-069-023-32
23	7	0.7	355	4	US-09-330-611-16
24	7	0.7	355	3	US-08-981-825-6
25	7	0.7	395	4	US-09-480-784-6
26	7	0.7	446	2	US-08-874-138-6
27	7	0.7	446	4	US-08-879-941-2

28	7	0.7	446	4	US-09-747-116-2	Sequence 2, Appl
29	7	0.7	456	4	US-09-134-001C-2853	Sequence 2853, Ap
30	7	0.7	474	4	US-09-538-414-8	Sequence 8, Appl
31	7	0.7	572	2	US-09-932-315-7	Sequence 7, Appl
32	7	0.7	572	2	US-08-993-318A-7	Sequence 7, Appl
33	7	0.7	572	4	US-09-399-886-7	Sequence 7, Appl
34	7	0.7	572	4	US-09-396-260-7	Sequence 7, Appl
35	7	0.7	572	4	US-09-576-281-7	Sequence 7, Appl
36	7	0.7	573	4	US-09-042-709A-19	Sequence 19, Appl
37	7	0.7	574	1	US-08-140-722A-7	Sequence 7, Appl
38	7	0.7	574	1	US-08-546-666-7	Sequence 7, Appl
39	7	0.7	574	2	US-08-916-745-7	Sequence 7, Appl
40	7	0.7	574	2	US-08-948-569A-4	Sequence 4, Appl
41	7	0.7	574	2	US-08-663-808-6	Sequence 6, Appl
42	7	0.7	574	2	US-09-042-929-7	Sequence 7, Appl
43	7	0.7	574	2	US-08-546-661-7	Sequence 7, Appl
44	7	0.7	574	2	US-09-042-960-7	Sequence 7, Appl
45	7	0.7	574	2	US-09-188-469-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-608-241-4
; Sequence 4, Application US/08608241
; Patent No. 5747328
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Witthuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,241
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 960296.93511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-9166
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-608-241-4

Query Match 0.8%; Score 8; DB 1; Length 224;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 418 PVTWVSSL 425
|||||||
DB 83 PVTWVSSL 90

RESULT 2
US-08-922-182-4
; Sequence 4, Application US/08922182
; Patent No. 5834300
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Withuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/922,182
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/608,241
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 960296.93511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-922-182-4
Query Match 0.8%; Score 8; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 418 PVVVSSL 425
Db 83 PVVVSSL 90
RESULT 3
US-08-919-953-4
; Sequence 4, Application US/08919953
; Patent No. 5837481
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Withuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,953
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/608,241
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 960296.93511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-919-953-4
Query Match 0.8%; Score 8; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 418 PVVVSSL 425
Db 83 PVVVSSL 90
RESULT 4
US-09-192-983-4
; Sequence 4, Application US/09192983A
; Patent No. 6242244
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy
; APPLICANT: Barber, Robert
; APPLICANT: Withuhn, Vernon
; TITLE OF INVENTION: Microbial System for Formaldehyde Sensing and
; FILE REFERENCE: 960296.95505
; CURRENT APPLICATION NUMBER: US/09/192,983A
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/919,953
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 08/608,241
; EARLIER FILING DATE: 1996-02-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Rhododactar sphaeroides
; US-09-192-983-4
Query Match 0.8%; Score 8; DB 4; Length 226;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 418 PVVVSSL 425
Db 85 PVVVSSL 92
RESULT 5
US-09-308-003-12
; Sequence 12, Application US/09308003

Patent No. 6326170
GENERAL INFORMATION:
APPLICANT: Burnham, Martin K. R.
APPLICANT: Lonetto, Michael A.
APPLICANT: Warren, Patrick V.
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
FILE REFERENCE: GMI10093
CURRENT APPLICATION NUMBER: US/09/308,003
CURRENT FILING DATE: 1999-05-10
EARLIER APPLICATION NUMBER: 60/058,710
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 390
TYPE: PR1
ORGANISM: Staphylococcus aureus
US-09-308-003-12

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 390;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 870 LPLVSVSL 877
|||||
Db 312 LPLVSVSL 319

RESULT 6
US-09-134-001C-4580
Sequence 4580, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4580
LENGTH: 396
TYPE: PR1
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4580

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 396;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 870 LPLVSVSL 877
|||||
Db 317 LPLVSVSL 324

RESULT 7
US-08-660-789-9
Sequence 9, Application US/08660789
Patent No. 5843405
GENERAL INFORMATION:
APPLICANT: Middeldorp, Jaap M.
TITLE OF INVENTION: Epstein Barr Virus peptides and
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 5843405el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville

STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,789
FILING DATE: 06-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
TELEFAX: (301) 977-0847
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Epstein-Barr virus
US-08-660-789-9

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 324 ASQASAG 330
|||||
Db 15 ASQASAG 21

RESULT 8
US-09-074-114-9
Sequence 9, Application US/09074114
Patent No. 6143865
GENERAL INFORMATION:
APPLICANT: Middeldorp, Jaap Michel
TITLE OF INVENTION: Epstein Barr Virus peptides and
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 6143865el Patent Dept.
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/074,114
FILING DATE: 07-MAY-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Klesner, Sharon N.
REGISTRATION NUMBER: 36,335
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-948-7400
TELEFAX: 301-948-9751
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Epstein-Barr virus
US-09-074-114-9

Query Match 0.7%; Score 7; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 ASOASAG 330
Db 15 ASOASAG 21

RESULT 9
US-08-127-351-8
Sequence 8, Application US/08127351
Patent No. 5449761

GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
TITLE OF INVENTION: CONSTRUCTS
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
ADDRESSEE: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,351
FILING DATE: 28-SEP-1993
CLASSIFICATION: 534
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide

US-08-127-351-8

Query Match 0.7%; Score 7; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LVDAALQ 76
Db 10 LVDAALQ 16

RESULT 10
US-08-480-367B-8
Sequence 8, Application US/08480367B

Patent No. 5578288
GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
TITLE OF INVENTION: CONSTRUCTS
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
STREET: 99 Canal Center Plaza, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,367B
FILING DATE: 07-06-95
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 2654-002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 684-1111
TELEFAX: (703) 684-1124
TELEX:

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide

US-08-480-367B-8

Query Match 0.7%; Score 7; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LVDAALQ 76
Db 10 LVDAALQ 16

RESULT 11
US-08-487-221A-8
Sequence 8, Application US/08487221A
Patent No. 5593656

GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
TITLE OF INVENTION: CONSTRUCTS
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
ADDRESSEE: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487, 221A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/127, 351
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-487-221A-8

Query Match 0.7%; Score 7; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LVDAALQ 76
|||||
DB 10 LVDAALQ 16

RESULT 12
US-08-480-370-8
Sequence 8, Application US/08480370
Patent No. 5609847
GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
NUMBER OF SEQUENCES: 56
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127, 351
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-127-351-9

TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-370-8

Query Match 0.7%; Score 7; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LVDAALQ 76
|||||
DB 10 LVDAALQ 16

RESULT 13
US-08-127-351-9
Sequence 9, Application US/08127351
Patent No. 5449761
GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
NUMBER OF SEQUENCES: 56
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127, 351
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-127-351-9

Query Match 0.7%; Score 7; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LVDAALQ 76
|||||
DB 10 LVDAALQ 16

RESULT 14
US-08-480-367B-9
; Sequence 9, Application US/08480367B
; Patent No. 5578288
; GENERAL INFORMATION:
; APPLICANT: BELINKA JR, BENJAMIN A.
; APPLICANT: COUGHLIN, DANIEL J.
; APPLICANT: ALVAREZ, VERNON L.
; APPLICANT: WOOD, RICHARD
; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
; TITLE OF INVENTION: CONSTRUCTS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,367B
; FILING DATE: 07-06-95
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Villacorta, Gilberto M.
; REGISTRATION NUMBER: 34,038
; REFERENCE/DOCKET NUMBER: 2654-002A
; TELEPHONE: (703) 684-1111
; TELEFAX: (703) 684-1124
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-480-367B-9

Query Match 0.7%; Score 7; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LVDAALQ 76
Db 10 LVDAALQ 16

RESULT 15
US-08-487-221A-9
; Sequence 9, Application US/08487221A
; Patent No. 5593656
; GENERAL INFORMATION:
; APPLICANT: BELINKA JR, BENJAMIN A.
; APPLICANT: COUGHLIN, DANIEL J.
; APPLICANT: ALVAREZ, VERNON L.
; APPLICANT: WOOD, RICHARD
; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
; TITLE OF INVENTION: CONSTRUCTS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER &
; ADDRESS: NEUSTADT, P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.

ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,221A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/127,351
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Villacorta, Gilberto M.
; REGISTRATION NUMBER: 34,038
; REFERENCE/DOCKET NUMBER: 4980-004-44
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-487-221A-9

Query Match 0.7%; Score 7; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LVDAALQ 76
Db 10 LVDAALQ 16

Search completed: May 7, 2003, 19:46:46
Job time: 22 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:44:39 ; Search time 29 Seconds
(without alignments)
3265.326 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 1029
Sequence: 1 MHILVHVAMVILTLIGPPRA.....MDPSAQISEALRIILHMEAVM 1029

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 349150 seqs, 92025710 residues

Word size: 0

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Published.Applications_AA.*

1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1029	100.0	1029	US-10-033-245-22	Sequence 22, Appl
2	1029	100.0	1029	US-10-033-223-22	Sequence 22, Appl
3	1029	100.0	1029	US-10-033-167-22	Sequence 22, Appl
4	1029	100.0	1029	US-10-033-244-22	Sequence 22, Appl
5	1029	100.0	1029	US-10-033-435-22	Sequence 22, Appl
6	1029	100.0	1029	US-10-033-990-22	Sequence 22, Appl
7	1029	100.0	1029	US-09-929-769-7	Sequence 7, Appl
8	1029	100.0	1029	US-10-033-996-22	Sequence 22, Appl
9	1029	100.0	1029	US-10-033-396-22	Sequence 22, Appl
10	1029	100.0	1029	US-10-033-246-22	Sequence 22, Appl
11	1029	100.0	1029	US-10-033-301-22	Sequence 22, Appl
12	1029	100.0	1029	US-10-033-326-22	Sequence 22, Appl
13	404	39.3	404	US-09-984-271-235	Sequence 235, App
14	327	31.8	328	US-09-984-271-148	Sequence 148, App
15	43	4.2	161	US-09-764-891-4155	Sequence 4155, App
16	8	0.8	209	US-09-893-737-126	Sequence 126, App
17	8	0.8	253	US-09-774-381-54	Sequence 54, App
18	7	0.7	9	US-09-922-226-88	Sequence 88, Appl
19	7	0.7	27	US-09-864-761-46828	Sequence 46828, A

20	7	0.7	54	10	US-09-864-761-46415	Sequence 46415, A
21	7	0.7	60	10	US-09-864-761-37408	Sequence 37408, A
22	7	0.7	62	9	US-09-764-868-1132	Sequence 1132, App
23	7	0.7	73	9	US-09-981-876-139	Sequence 139, App
24	7	0.7	73	9	US-09-148-545-139	Sequence 139, App
25	7	0.7	95	9	US-10-016-634A-125	Sequence 125, App
26	7	0.7	99	10	US-09-864-761-46653	Sequence 46653, A
27	7	0.7	116	10	US-09-864-761-44770	Sequence 44770, A
28	7	0.7	172	9	US-09-847-208-50	Sequence 50, Appl
29	7	0.7	181	9	US-10-163-499-14	Sequence 14, Appl
30	7	0.7	181	9	US-10-163-499-16	Sequence 16, Appl
31	7	0.7	197	9	US-10-097-065-135	Sequence 135, App
32	7	0.7	208	10	US-09-214-881A-11	Sequence 11, Appl
33	7	0.7	233	10	US-09-938-803-15	Sequence 15, Appl
34	7	0.7	262	10	US-09-960-738A-2	Sequence 2, Appl
35	7	0.7	271	10	US-09-846-808-20	Sequence 20, Appl
36	7	0.7	276	9	US-09-738-626-5599	Sequence 5599, App
37	7	0.7	281	9	US-10-101-464A-518	Sequence 518, App
38	7	0.7	283	10	US-09-895-913A-28	Sequence 28, Appl
39	7	0.7	283	10	US-09-815-242-14063	Sequence 14063, A
40	7	0.7	286	9	US-09-997-816-2	Sequence 2, Appl
41	7	0.7	298	10	US-09-815-242-5142	Sequence 5142, App
42	7	0.7	315	10	US-09-741-669-361	Sequence 361, App
43	7	0.7	319	9	US-10-116-722A-6	Sequence 6, Appl
44	7	0.7	319	9	US-10-116-722A-8	Sequence 8, Appl
45	7	0.7	320	10	US-09-741-669-467	Sequence 467, App

ALIGNMENTS

RESULT 1
US-10-033-245-22
Sequence 22, Application US/10033245
Patent No. US20020160392A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gunney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1C7
CURRENT APPLICATION NUMBER: US/10/033,245
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965

PRIOR FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: 60/162,506
 PRIOR FILING DATE: 1999-10-29
 PRIOR APPLICATION NUMBER: 60/170,262
 PRIOR FILING DATE: 1999-12-09
 PRIOR APPLICATION NUMBER: 60/187,202
 PRIOR FILING DATE: 2000-03-03
 PRIOR APPLICATION NUMBER: PCT/US99/12252
 PRIOR FILING DATE: 1999-06-02
 PRIOR APPLICATION NUMBER: PCT/US99/28634
 PRIOR FILING DATE: 1999-12-01
 PRIOR APPLICATION NUMBER: PCT/US99/28551
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US00/03565
 PRIOR FILING DATE: 2000-02-11
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: PCT/US00/05841
 PRIOR FILING DATE: 2000-03-02
 PRIOR APPLICATION NUMBER: PCT/US00/08439
 PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: PCT/US00/14941
 PRIOR FILING DATE: 2000-05-30
 PRIOR APPLICATION NUMBER: PCT/US00/15264
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: 2000-12-01
 NUMBER OF SEQ ID NOS: 38
 SEQ ID NO 22
 LENGTH: 1029
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-033-245-22

Query Match 100.0%; Score 1029; DB 9; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHILVNAVMTLTGPRRADDSEFOALLDIPPEKPLPAFLVDTSEKLLPDMKL 60
 DB 1 MHILVNAVMTLTGPRRADDSEFOALLDIPPEKPLPAFLVDTSEKLLPDMKL 60
 QY 61 RMIRSEVLRLVDAALQDLEPOQLLFVQSGFIPVSSMSKLLQFDQVAHDPOTLEONIM 120
 DB 61 RMIRSEVLRLVDAALQDLEPOQLLFVQSGFIPVSSMSKLLQFDQVAHDPOTLEONIM 120
 QY 121 DKNYMAHLVEVQHERGASGGOTFHSLLTASLPPRRDSTEAPKPSSEOPIGGRIRVGT 180
 DB 121 DKNYMAHLVEVQHERGASGGOTFHSLLTASLPPRRDSTEAPKPSSEOPIGGRIRVGT 180
 QY 181 QLRVLGEDDLACGFLQIFPLSPRPMWSSPRVALALQALQELARVVOGSEVPVGI 240
 DB 181 QLRVLGEDDLACGFLQIFPLSPRPMWSSPRVALALQALQELARVVOGSEVPVGI 240
 QY 241 TVRVLQALATLTLSSPHGALVMSHRSHPFLACPLLRQLOYORCVPDGTGSSLFLAVLL 300
 DB 241 TVRVLQALATLTLSSPHGALVMSHRSHPFLACPLLRQLOYORCVPDGTGSSLFLAVLL 300
 QY 301 QMLQWLDSPEVEGGPLRAQDRLMLASQASAGRRLSDVVGGLRLAEALAFRODLEVVSTV 360
 DB 301 QMLQWLDSPEVEGGPLRAQDRLMLASQASAGRRLSDVVGGLRLAEALAFRODLEVVSTV 360
 QY 361 RAVIATIRSEQCSVEPDLISKVQLIEVRSPLHEELLTAFFSATADAASPPACKPVV 420
 DB 361 RAVIATIRSEQCSVEPDLISKVQLIEVRSPLHEELLTAFFSATADAASPPACKPVV 420
 QY 421 VVSSLIOEEEPGLAGRGAGDGSLEAVRLGSPSSGLLVDMLEMDPEVSSCPDLQRL 480
 DB 421 VVSSLIOEEEPGLAGRGAGDGSLEAVRLGSPSSGLLVDMLEMDPEVSSCPDLQRL 480
 QY 481 FSRKRGKQAVPSFRYLLTLFTTHOSSMPTLHQCIRVLLGKRSRECFDSSASLDFLMAC 540
 DB 481 FSRKRGKQAVPSFRYLLTLFTTHOSSMPTLHQCIRVLLGKRSRECFDSSASLDFLMAC 540

QY 541 IHVERIMQGRDQRTPOKRRRELYLRVQGPGLISLVELILAEATRSODGDTAACSLIOAR 600
 DB 541 IHVERIMQGRDQRTPOKRRRELYLRVQGPGLISLVELILAEATRSODGDTAACSLIOAR 600
 QY 601 LPILLSCCGDDESVKRYTEHLSCIOQMGDSVIGRCRDLLOLYQREPRLPVPEVL 660
 DB 601 LPILLSCCGDDESVKRYTEHLSCIOQMGDSVIGRCRDLLOLYQREPRLPVPEVL 660
 QY 661 LHSEGAASSVCKIDGLIHRTTLADTSDSRALENGADASMACRLAAVAPLLLRHL 720
 DB 661 LHSEGAASSVCKIDGLIHRTTLADTSDSRALENGADASMACRLAAVAPLLLRHL 720
 QY 721 PMIALHGRTHLNFQFROONHLSCLHVLGLLELLQPHVFRSEHOGALMDCLSLFRL 780
 DB 721 PMIALHGRTHLNFQFROONHLSCLHVLGLLELLQPHVFRSEHOGALMDCLSLFRL 780
 QY 781 LLNKRSSRHAAFINKEVOFIHKYIYVNAAPASIFLOKHADPLHDSFONSDDLVMKSL 840
 DB 781 LLNKRSSRHAAFINKEVOFIHKYIYVNAAPASIFLOKHADPLHDSFONSDDLVMKSL 840
 QY 841 LAGLSRDRDRDRLGDEEGEESAGSLPLVSVSLFTPLTAEMAPYKKRLSRGOTVE 900
 DB 841 LAGLSRDRDRDRLGDEEGEESAGSLPLVSVSLFTPLTAEMAPYKKRLSRGOTVE 900
 QY 901 DLEVLSDIDEMSRRELLISFSTNLQRLMSSAEECCRLNLAISLRMONSPSTIAAF 960
 DB 901 DLEVLSDIDEMSRRELLISFSTNLQRLMSSAEECCRLNLAISLRMONSPSTIAAF 960
 QY 961 LPTMYCLGSODPEVOTARLNLPEVALLOEHAVALHRAFLVGMVGMOPSAQISEAL 1020
 DB 961 LPTMYCLGSODPEVOTARLNLPEVALLOEHAVALHRAFLVGMVGMOPSAQISEAL 1020
 QY 1021 RLHMEAVM 1029
 DB 1021 RLHMEAVM 1029

RESULT 2
 US-10-033-223-22
 ; Sequence 22, Application US/10033223
 ; Patent No. US20020164646A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2930R1C9
 ; CURRENT APPLICATION NUMBER: US/10/033,223
 ; PRIOR FILING DATE: 2001-12-27
 ; PRIOR APPLICATION NUMBER: 60/095,325
 ; PRIOR FILING DATE: 1998-08-04
 ; PRIOR APPLICATION NUMBER: 60/112,851
 ; PRIOR FILING DATE: 1998-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,145
 ; PRIOR FILING DATE: 1998-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,511
 ; PRIOR FILING DATE: 1998-12-22
 ; PRIOR APPLICATION NUMBER: 60/115,558
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/115,565
 ; PRIOR FILING DATE: 1999-01-12

```

: PRIOR APPLICATION NUMBER: 60/115,733
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/119,341
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: 60/119,537
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: 60/119,965
: PRIOR FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: 60/162,506
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: 60/170,262
: PRIOR FILING DATE: 1999-12-09
: PRIOR APPLICATION NUMBER: 60/187,202
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: PCT/US99/28634
: PRIOR FILING DATE: 1999-12-01
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US00/03565
: PRIOR FILING DATE: 2000-02-11
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: 2000-03-02
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: PCT/US00/14941
: PRIOR FILING DATE: 2000-05-30
: PRIOR APPLICATION NUMBER: PCT/US00/15264
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: 2000-12-01
: NUMBER OF SEQ ID NOS: 38
: SEQ ID NO 22
: LENGTH: 1029
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-033-223-22

```

```

Query Match      100.0%; Score 1029; DB 9; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MHILVHAMVILTLGPPRADSEFOALDIFPEEKPLPTAFIVDTSEALLDPWKL 60
DB 1 MHILVHAMVILTLGPPRADSEFOALDIFPEEKPLPTAFIVDTSEALLDPWKL 60
QY 61 RMIRSEVLRLVDALQDLEPQQLLFVQSGFIPVSSMSKILQFLDOAVAHDPQILEONIM 120
DB 61 RMIRSEVLRLVDALQDLEPQQLLFVQSGFIPVSSMSKILQFLDOAVAHDPQILEONIM 120
QY 121 DKNYMAHLVEVOHERGASGGQTFHSLTLASLPDRDSTEARPKSSPEQPIQGRIRVGT 180
DB 121 DKNYMAHLVEVOHERGASGGQTFHSLTLASLPDRDSTEARPKSSPEQPIQGRIRVGT 180
QY 181 QLRVIGPEDDLAGMFLQIFPLSPDRMOSSSRPVALAQAQAGELAVVGGSPVPPI 240
DB 181 QLRVIGPEDDLAGMFLQIFPLSPDRMOSSSRPVALAQAQAGELAVVGGSPVPPI 240
QY 241 TVRVLOALATLTLSSPHGALVMSMRSHFLACPLRLQOLCOYORCVPODTGFSSLEFLKVL 300
DB 241 TVRVLOALATLTLSSPHGALVMSMRSHFLACPLRLQOLCOYORCVPODTGFSSLEFLKVL 300
QY 301 QMLQMLDSPEGVGGPLRLQRLMLASQASAGRRLSDVRGGLRLAALAFRODLEVVSTIV 360
DB 301 QMLQMLDSPEGVGGPLRLQRLMLASQASAGRRLSDVRGGLRLAALAFRODLEVVSTIV 360
QY 361 RAVITFLSGECCSVPEPDLISKVLQGLIEVRSPHLELLTAFSAFADAASFPACKPV 420
DB 361 RAVITFLSGECCSVPEPDLISKVLQGLIEVRSPHLELLTAFSAFADAASFPACKPV 420

```

```

QY 421 VVSSILLQEEEPBLAGKPGADGSLAVNRLGPSSGLVDMLEMLDPVEVSSCPDQLRL 480
DB 421 VVSSILLQEEEPBLAGKPGADGSLAVNRLGPSSGLVDMLEMLDPVEVSSCPDQLRL 480
QY 481 FSRKRGKGAQVSPRPYLLTLTHQSSWPTLHQCIRVLLGKSRQOREPDSASLDFLMAC 540
DB 481 FSRKRGKGAQVSPRPYLLTLTHQSSWPTLHQCIRVLLGKSRQOREPDSASLDFLMAC 540
QY 541 HVPRWMOGRDORTQKRRREELVLRVQGPBELLISVELLILAEETNSQDDTAAACSLIQAR 600
DB 541 HVPRWMOGRDORTQKRRREELVLRVQGPBELLISVELLILAEETNSQDDTAAACSLIQAR 600
QY 601 LPDLLSCCGDDESVRYKTEHLSGCIQMGDSVLGRRCBDLLQLYLQPRELVRPEVL 660
DB 601 LPDLLSCCGDDESVRYKTEHLSGCIQMGDSVLGRRCBDLLQLYLQPRELVRPEVL 660
QY 661 LHSEGAASSVCKLDGLHRTTLADTSDSRALENRGADASMACRKLAVAHPLLLRL 720
DB 661 LHSEGAASSVCKLDGLHRTTLADTSDSRALENRGADASMACRKLAVAHPLLLRL 720
QY 721 PMIAALHGRTHLNFQERQONHLSCFLHVLGLLELQPHVRSRHOGLMPCLLSFTPL 780
DB 721 PMIAALHGRTHLNFQERQONHLSCFLHVLGLLELQPHVRSRHOGLMPCLLSFTPL 780
QY 781 LLNRYKSSRHIAFINKFVQFIHKYTTYNAPAISEFLQKHADPLHDLSFDSNDLYWLKSL 840
DB 781 LLNRYKSSRHIAFINKFVQFIHKYTTYNAPAISEFLQKHADPLHDLSFDSNDLYWLKSL 840
QY 841 LAGLSLPSRDRTDRLGDEEGEESAGSLPVYSLEFPLTAEMAPYMKRLSKQIVE 900
DB 841 LAGLSLPSRDRTDRLGDEEGEESAGSLPVYSLEFPLTAEMAPYMKRLSKQIVE 900
QY 901 DLEVLSDIDEMSRREPILTSFSTMLQRLMSAECCCNLAFLSALRSMQSPSTAAAF 960
DB 901 DLEVLSDIDEMSRREPILTSFSTMLQRLMSAECCCNLAFLSALRSMQSPSTAAAF 960
QY 961 LPTFMVCLGSDPEFVQATRLNLPEVALLCQHAANVLHRAFLVGMYGMDPSAQISEAL 1020
DB 961 LPTFMVCLGSDPEFVQATRLNLPEVALLCQHAANVLHRAFLVGMYGMDPSAQISEAL 1020
QY 1021 RLHMEAVM 1029
DB 1021 RLHMEAVM 1029

```

```

RESULT 3
US-10-033-167-22
: Sequence 22, Application US/10033167
: Publication No. US20020182618A1
: GENERAL INFORMATION:
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2930RC10
: CURRENT APPLICATION NUMBER: US/10/033,167
: PRIOR FILING DATE: 2001-12-27
: PRIOR APPLICATION NUMBER: 60/095,325
: PRIOR FILING DATE: 1998-08-04
: PRIOR APPLICATION NUMBER: 60/112,851
: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,145

```

```

: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,511
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: 60/115,558
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/115,565
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/115,733
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/119,341
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: 60/119,537
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: 60/119,965
: PRIOR FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: 60/162,506
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: 60/170,262
: PRIOR FILING DATE: 1999-12-09
: PRIOR APPLICATION NUMBER: 60/187,202
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: PCT/US99/28634
: PRIOR FILING DATE: 1999-12-01
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US00/03565
: PRIOR FILING DATE: 2000-02-11
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: 2000-03-02
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: PCT/US00/14941
: PRIOR FILING DATE: 2000-05-30
: PRIOR APPLICATION NUMBER: PCT/US00/15264
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: 2000-12-01
: NUMBER OF SEQ ID NOS: 36
: SEQ ID NO 22
: LENGTH: 1029
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-033-167-22

```

Query Match 100.0% Score 1029: DB 9: Length 1029:
 Best Local Similarity 100.0%: Pred. No. 0:
 Matches 1029: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

```

QY 1 MHILVNAVYILTLGPRADSEFOALDIWPEERKPLPTAFVDTSEBALPDKL 60
DB 1 MHILVNAVYILTLGPRADSEFOALDIWPEERKPLPTAFVDTSEBALPDKL 60
QY 61 RMIRSEVLRVDAADLEPOQLLEFYQSGIYVSSSKLLOFLDOAVAHDPOTLEONIM 120
DB 61 RMIRSEVLRVDAADLEPOQLLEFYQSGIYVSSSKLLOFLDOAVAHDPOTLEONIM 120
QY 121 DKNYMAHLVVOHERGASGOTFHSILTLASLPPRODSTEAPKPKSSPEOPIGGIRIVGT 180
DB 121 DKNYMAHLVVOHERGASGOTFHSILTLASLPPRODSTEAPKPKSSPEOPIGGIRIVGT 180
QY 181 QLRVLGPEDDLAGMFLQIFPLSPDRPMWSSPPVALALQALGOELARVVOGSPFVG 240
DB 181 QLRVLGPEDDLAGMFLQIFPLSPDRPMWSSPPVALALQALGOELARVVOGSPFVG 240
QY 241 TVRYLQALATLSSPHGALVMSHNSHFLACPLRLQCOYRCVPDGTGSSFLFKVLL 300
DB 241 TVRYLQALATLSSPHGALVMSHNSHFLACPLRLQCOYRCVPDGTGSSFLFKVLL 300
QY 301 QMLQWIDSPVEGGLPLRAQLRMLASQASAGRRLSDVAGGLRLALAEALAFRODLEVVSSV 360

```

```

DB 301 QMLQWIDSPVEGGLPLRAQLRMLASQASAGRRLSDVAGGLRLALAEALAFRODLEVVSSV 360
QY 361 RAVIATLRSEGOCSVEPDLISKVLOGLIEVRSRHEBELLAFFSATADAAAPPACKPVY 420
DB 361 RAVIATLRSEGOCSVEPDLISKVLOGLIEVRSRHEBELLAFFSATADAAAPPACKPVY 420
QY 421 VVSSLLIQEEEPPLAGKRGADGSLAVRIGPSSGLVMDLEMLDEPVVSSCPDLQRL 480
DB 421 VVSSLLIQEEEPPLAGKRGADGSLAVRIGPSSGLVMDLEMLDEPVVSSCPDLQRL 480
QY 481 FSRKRGKGAQVPSFRYLLTLFTHOSWPTLHOCIRVLLGKSREOFDPDSASIDFLMAC 540
DB 481 FSRKRGKGAQVPSFRYLLTLFTHOSWPTLHOCIRVLLGKSREOFDPDSASIDFLMAC 540
QY 541 IHVPRIMQGRDQRTPOKREELVRYOGPELISVELILAEATRSODGDTAACSLIOAR 600
DB 541 IHVPRIMQGRDQRTPOKREELVRYOGPELISVELILAEATRSODGDTAACSLIOAR 600
QY 601 LPLILSCCGDDDESVRKVTETHLSGCTIOQWGDVYGRRCRLQLQYLQRELRVPPEVL 660
DB 601 LPLILSCCGDDDESVRKVTETHLSGCTIOQWGDVYGRRCRLQLQYLQRELRVPPEVL 660
QY 661 LHS6AASSSVCKLDGILHRFTLLADTSDSRALENRGADAMACRKLAAVHPULLLRL 720
DB 661 LHS6AASSSVCKLDGILHRFTLLADTSDSRALENRGADAMACRKLAAVHPULLLRL 720
QY 721 PMIAALLHGRTHLNFOEPRQOONHLSCLFHLVGLLELQHPVFRSEHOGALMDCLSLFRL 780
DB 721 PMIAALLHGRTHLNFOEPRQOONHLSCLFHLVGLLELQHPVFRSEHOGALMDCLSLFRL 780
QY 781 LMYRKSSRLAFLINKFVFIHKYTYNAPPAISFLQKADPHDLSFNSDLYMLKSL 840
DB 781 LMYRKSSRLAFLINKFVFIHKYTYNAPPAISFLQKADPHDLSFNSDLYMLKSL 840
QY 841 LAGLSLRDQDRDRCDEGEESAGSLPLVSVSLFTEUTLAEMA PYKKRLSRGOTVE 900
DB 841 LAGLSLRDQDRDRCDEGEESAGSLPLVSVSLFTEUTLAEMA PYKKRLSRGOTVE 900
QY 901 DLLEVLSDIDEMSRREILSFSTNLQRLMSSAECCRNLAFLSLRMONSPSIAAFL 960
DB 901 DLLEVLSDIDEMSRREILSFSTNLQRLMSSAECCRNLAFLSLRMONSPSIAAFL 960
QY 961 LPTFMVCLGSDQDEEVOTALRNLPEYALLCOEHAAVLLHRAFLVGMGQMDPSAOISEAL 1020
DB 961 LPTFMVCLGSDQDEEVOTALRNLPEYALLCOEHAAVLLHRAFLVGMGQMDPSAOISEAL 1020
QY 1021 RILHMEAVM 1029
DB 1021 RILHMEAVM 1029

```

RESULT 4
 US-10-033-244-22
 Publication NO. US20020192668A1
 GENERAL INFORMATION:
 APPLICANT: Bolstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Goddard, Audrey
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: P2930R1C2


```

: CURRENT APPLICATION NUMBER: US/10/033,244
: CURRENT FILING DATE: 2001-12-27
: PRIOR APPLICATION NUMBER: 60/095,325
: PRIOR FILING DATE: 1998-08-04
: PRIOR APPLICATION NUMBER: 60/112,851
: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,145
: PRIOR FILING DATE: 1998-12-16
: PRIOR APPLICATION NUMBER: 60/113,511
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: 60/115,558
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/115,565
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/115,733
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/119,341
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: 60/119,537
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: 60/119,965
: PRIOR FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: 60/162,506
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: 60/170,262
: PRIOR FILING DATE: 1999-12-09
: PRIOR APPLICATION NUMBER: 60/187,202
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: PCT/US99/28634
: PRIOR FILING DATE: 1999-12-01
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US00/03565
: PRIOR FILING DATE: 2000-02-11
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: 2000-03-02
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: PCT/US00/14941
: PRIOR FILING DATE: 2000-05-30
: PRIOR APPLICATION NUMBER: PCT/US00/15264
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: 2000-12-01
: NUMBER OF SEQ ID NOS: 38
: SEQ ID NO 22
: LENGTH: 1029
: TYPE: prt
: ORGANISM: Homo sapiens
US-10-033-244-22

Query Match      100.0%  Score 1029;  DB 9;  Length 1029;
Best Local Similarity 100.0%;  Pred. No.: 0;
Matches 1029;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Db 181 QLRVLCPEDDLGMFLQIFPLSPDRMOSSSPRVALALQALGQELARVVOGSEVPVGI 240
QY 241 TVRVQLATLTLSSPHGALVMSMRSHFLACPLRLQLOYQRCVQPDGFSSTLFLKVL 300
Db 241 TVRVQLATLTLSSPHGALVMSMRSHFLACPLRLQLOYQRCVQPDGFSSTLFLKVL 300
QY 301 QMLQWIDSPGVEGCPRLAOLRLMLASQASAGRRLSDVRGGLRLRLAEALAFRODLEVVSTV 360
Db 301 QMLQWIDSPGVEGCPRLAOLRLMLASQASAGRRLSDVRGGLRLRLAEALAFRODLEVVSTV 360
QY 361 RAVIATLRSGECCSVEPDLISKVLOGLIEVRSPhLEELLTAFFSATADAAPPACKPVV 420
Db 361 RAVIATLRSGECCSVEPDLISKVLOGLIEVRSPhLEELLTAFFSATADAAPPACKPVV 420
QY 421 VVSILLQEEEPPLAGKPPADGSLAVRLGSSGLVWMLDPEVVSQCPDOLRL 480
Db 421 VVSILLQEEEPPLAGKPPADGSLAVRLGSSGLVWMLDPEVVSQCPDOLRL 480
QY 481 FSRKKGQAOVPSFRPYLTTLTTHOSSWPTLHQCIRVLGKSREQRPFSASLDPLMAC 540
Db 481 FSRKKGQAOVPSFRPYLTTLTTHOSSWPTLHQCIRVLGKSREQRPFSASLDPLMAC 540
QY 541 IHVPRIWGRDORTPOKRREELVLRVQPELISVLELLIAEAFTRSQDDDTAACSLIQAR 600
Db 541 IHVPRIWGRDORTPOKRREELVLRVQPELISVLELLIAEAFTRSQDDDTAACSLIQAR 600
QY 601 LPULLSCCGDDESVKRVTEHLSGCTIQWGDVSLGRCDLLQLTLQRPRLVVPVYL 660
Db 601 LPULLSCCGDDESVKRVTEHLSGCTIQWGDVSLGRCDLLQLTLQRPRLVVPVYL 660
QY 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRKLAVAPHLRL 720
Db 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRKLAVAPHLRL 720
QY 721 PMIAALHRTLNROEFQOQNLSCFLVLYGLLELQHVHRSSEQALMDCLSFTL 780
Db 721 PMIAALHRTLNROEFQOQNLSCFLVLYGLLELQHVHRSSEQALMDCLSFTL 780
QY 781 LNTYKSSSHLAAFIKFFQFIHKYTYNAPAAISFLQKHADPHLSFSDNSDLVWLKSL 840
Db 781 LNTYKSSSHLAAFIKFFQFIHKYTYNAPAAISFLQKHADPHLSFSDNSDLVWLKSL 840
QY 841 LAGLSLPSRDDRTDGLDEGESESSAGSLPLVSVSLFTPLTAEMAPYMKRLSRQOVE 900
Db 841 LAGLSLPSRDDRTDGLDEGESESSAGSLPLVSVSLFTPLTAEMAPYMKRLSRQOVE 900
QY 901 DLEVLSDIDEMSRRRPELTSFSTNLQRLMSAECCCNLAFAFLRLRMQNSPSTAAAF 960
Db 901 DLEVLSDIDEMSRRRPELTSFSTNLQRLMSAECCCNLAFAFLRLRMQNSPSTAAAF 960
QY 961 LPTFMYCISQDPFEVYQALRLNPEYALICQEHAAVLLHRAFLVGYGQMDPSAQISEAL 1020
Db 961 LPTFMYCISQDPFEVYQALRLNPEYALICQEHAAVLLHRAFLVGYGQMDPSAQISEAL 1020
QY 1021 RILHMEAVM 1029
Db 1021 RILHMEAVM 1029

RESULT 5
US-10-033-435-22
: Sequence 22, Application US/10033435
: Publication No. US20030027256A1
: GENERAL INFORMATION:
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Roy, Margaret Ann

```

```

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1CS
CURRENT APPLICATION NUMBER: US/10/033,435
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
SEQUENCE ID NO 22
SEQUENCE ID NOS: 38
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-10-033-435-22
Query Match 100.0%; Score 1029; DB 9; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

121 DKNYMAHLVEYQHEBRCASGCGTFHSLTASLPDRDSTEAPKPSSEOPIGGRIRVGT 180
121 DKNYMAHLVEYQHEBRCASGCGTFHSLTASLPDRDSTEAPKPSSEOPIGGRIRVGT 180
181 QLRVLPEDDLAGMFLQIFPLSPDRMOSSPRVALALQALQGLARVYQSGPEVPGI 240
181 QLRVLPEDDLAGMFLQIFPLSPDRMOSSPRVALALQALQGLARVYQSGPEVPGI 240
241 TVRVQLATLTLSSPHGALVMSHRSHPACPLRLQCOVQRPVDPDTGSSFLVLVLL 300
241 TVRVQLATLTLSSPHGALVMSHRSHPACPLRLQCOVQRPVDPDTGSSFLVLVLL 300
301 QMLQMLDPSGVEGGPLRAQRLMLASQASAGRLSDVAGLRLALAFRODLEVSSTV 360
301 QMLQMLDPSGVEGGPLRAQRLMLASQASAGRLSDVAGLRLALAFRODLEVSSTV 360
361 RAVIATLRSGEGCSVEEDLISKVLQGLIEVRSPLHEELTAFFSATADAAFPACKPVV 420
361 RAVIATLRSGEGCSVEEDLISKVLQGLIEVRSPLHEELTAFFSATADAAFPACKPVV 420
421 VVSSLLOEPEEPLAGKPGADGSLAEAVRGLVMDLMDPEVSSCPDLORL 480
421 VVSSLLOEPEEPLAGKPGADGSLAEAVRGLVMDLMDPEVSSCPDLORL 480
481 FSRKRGQAQVPSFRPYLLTLFTHQSSWPTLMOICIRVLLGKSRQRPDPSASIDFLMAC 540
481 FSRKRGQAQVPSFRPYLLTLFTHQSSWPTLMOICIRVLLGKSRQRPDPSASIDFLMAC 540
541 IHVRINQGRQRTPOKRRRELVLRVQGPGLISVELLAEAEFRSDGDGAASLQAR 600
541 IHVRINQGRQRTPOKRRRELVLRVQGPGLISVELLAEAEFRSDGDGAASLQAR 600
601 LPILLSCCGDDESVRRVTEHLSGCIQWGDVYGRRCRDLLQDLYQRELRVPEVL 660
601 LPILLSCCGDDESVRRVTEHLSGCIQWGDVYGRRCRDLLQDLYQRELRVPEVL 660
661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRKLAAVHPILLRL 720
661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRKLAAVHPILLRL 720
721 PMTALLHGRTHLNFQEFROONHLSCLFVHVLGELLELOPHVREHOGALMDCLSTIRL 780
721 PMTALLHGRTHLNFQEFROONHLSCLFVHVLGELLELOPHVREHOGALMDCLSTIRL 780
781 LLYNRKSSRHIAFINKFVOFIHKYITYNAPAAISFLQKHADPLHDSFPNSDLVMLKSL 840
781 LLYNRKSSRHIAFINKFVOFIHKYITYNAPAAISFLQKHADPLHDSFPNSDLVMLKSL 840
841 LAGLSLPSRDDRDRLGDEBGEESAGSLPLVSVSLFTPLTAEMAAPYKRLSRGQTV 900
841 LAGLSLPSRDDRDRLGDEBGEESAGSLPLVSVSLFTPLTAEMAAPYKRLSRGQTV 900
901 DLEVLSDIDEMSRREILISFSTNLQRLMSSAECCRNLAFLARSMONSSTIAAF 960
901 DLEVLSDIDEMSRREILISFSTNLQRLMSSAECCRNLAFLARSMONSSTIAAF 960
961 LPTFMVCLGSODEEVQVOTALRNLPYVALLCOEHAVALLHRAFLVGMVGMQDPASQISEAL 1020
961 LPTFMVCLGSODEEVQVOTALRNLPYVALLCOEHAVALLHRAFLVGMVGMQDPASQISEAL 1020
1021 RILHMEAVM 1029
1021 RILHMEAVM 1029
US-10-032-990-22
Sequence 22, Application US/10032990
Publication No. US20030032060A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc

```

APPLICANT: Ferrara, Napoleone
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Goddard, Audrey
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tamas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P2930R1C11
 CURRENT FILING DATE: 2001-12-27
 PRIOR APPLICATION NUMBER: 60/095,325
 PRIOR FILING DATE: 1998-08-04
 PRIOR APPLICATION NUMBER: 60/112,851
 PRIOR FILING DATE: 1998-12-16
 PRIOR APPLICATION NUMBER: 60/113,145
 PRIOR FILING DATE: 1998-12-16
 PRIOR APPLICATION NUMBER: 60/113,511
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/115,558
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115,565
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115,733
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/119,341
 PRIOR FILING DATE: 1999-02-09
 PRIOR APPLICATION NUMBER: 60/119,537
 PRIOR FILING DATE: 1999-02-10
 PRIOR APPLICATION NUMBER: 60/119,965
 PRIOR FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: 60/162,506
 PRIOR FILING DATE: 1999-10-29
 PRIOR APPLICATION NUMBER: 60/170,262
 PRIOR FILING DATE: 1999-12-09
 PRIOR APPLICATION NUMBER: 60/187,202
 PRIOR FILING DATE: 2000-03-03
 PRIOR APPLICATION NUMBER: PCT/US99/12252
 PRIOR FILING DATE: 1999-06-02
 PRIOR APPLICATION NUMBER: PCT/US99/28634
 PRIOR FILING DATE: 1999-12-01
 PRIOR APPLICATION NUMBER: PCT/US99/28551
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US00/03565
 PRIOR FILING DATE: 2000-02-11
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: PCT/US00/05841
 PRIOR FILING DATE: 2000-03-02
 PRIOR APPLICATION NUMBER: PCT/US00/08439
 PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: PCT/US00/14941
 PRIOR FILING DATE: 2000-05-30
 PRIOR APPLICATION NUMBER: PCT/US00/15264
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 NUMBER OF SEQ ID NOS: 38
 SEQ ID NO 22
 LENGTH: 1029
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-032-990-22

Query Match 100.0%; Score 1029; DB 9; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MHILVVHNAVYILLTLGPPRADSEFOALDIMEPEEKPLPTAFVYDTSEALLPDMKL	60
DB	1	MHILVVHNAVYILLTLGPPRADSEFOALDIMEPEEKPLPTAFVYDTSEALLPDMKL	60
QY	61	RMIRSEVLRVDAIADDEPQOLLFVSGFIPVSSMSKLPLODAVAHDPQTEQNIM	120
DB	61	RMIRSEVLRVDAIADDEPQOLLFVSGFIPVSSMSKLPLODAVAHDPQTEQNIM	120
QY	121	DKNYMAHLYEVOHERGASGGOTFHSILTASLPBRSTAPKPSPEPIGGRTIRVGT	180
DB	121	DKNYMAHLYEVOHERGASGGOTFHSILTASLPBRSTAPKPSPEPIGGRTIRVGT	180
QY	181	OLRVLGPEDDLAGMFLQIPIPLSPDPWQSSPPRVALAQALGQELAVVQSSPEVPGI	240
DB	181	OLRVLGPEDDLAGMFLQIPIPLSPDPWQSSPPRVALAQALGQELAVVQSSPEVPGI	240
QY	241	TVRVYQALATLLSSPHGALVMSMRSHTLACPRLQCLQYQRCVQDTGFSSFLKVL	300
DB	241	TVRVYQALATLLSSPHGALVMSMRSHTLACPRLQCLQYQRCVQDTGFSSFLKVL	300
QY	301	OMLQWMDSPGVGGPLRAQLRMLASQASAGRISDVRGGLRLAEALARQDLEVVSTV	360
DB	301	OMLQWMDSPGVGGPLRAQLRMLASQASAGRISDVRGGLRLAEALARQDLEVVSTV	360
QY	361	RAVIATLRSGECSVEPDLISKVLQGLIEVRSFHEBELLTAFPSATADAAPFPACKPVY	420
DB	361	RAVIATLRSGECSVEPDLISKVLQGLIEVRSFHEBELLTAFPSATADAAPFPACKPVY	420
QY	421	VVSSLLOEPEPLAGKPGADGSLAVRLGSSGLVWLEMLDEPVVSSCPDLORL	480
DB	421	VVSSLLOEPEPLAGKPGADGSLAVRLGSSGLVWLEMLDEPVVSSCPDLORL	480
QY	481	FSRRKKGQAQVPSFRPYLLTFTHOSSMPTLHQCIRVLLGSRREPRPSASIDPLMC	540
DB	481	FSRRKKGQAQVPSFRPYLLTFTHOSSMPTLHQCIRVLLGSRREPRPSASIDPLMC	540
QY	541	HNVPRIWGRDQRTPOKRREELVLRVQPELISVLELLAEATRSQDDDTAACSILQAR	600
DB	541	HNVPRIWGRDQRTPOKRREELVLRVQPELISVLELLAEATRSQDDDTAACSILQAR	600
QY	601	LPILLSCCGGDESVKVTENHSGCTIQMGDSVYLRRCRDLLQLYLQREPLRVPEYL	660
DB	601	LPILLSCCGGDESVKVTENHSGCTIQMGDSVYLRRCRDLLQLYLQREPLRVPEYL	660
QY	661	LHSEGAASSVCKLQGLIHRFTLLADTSDSRALENRGADASACKRLVAHPDLLLRHL	720
DB	661	LHSEGAASSVCKLQGLIHRFTLLADTSDSRALENRGADASACKRLVAHPDLLLRHL	720
QY	721	PMIALHGRTHLNFQEFROQNHLSCELVHGLLELLQPHVFRSEHOGALMDCILSFIRL	780
DB	721	PMIALHGRTHLNFQEFROQNHLSCELVHGLLELLQPHVFRSEHOGALMDCILSFIRL	780
QY	781	LNLYKSSRHLLAFINKFVQFIHKYTTYNAPAAISFLQKHADPLHDLSFSDNSDLVMLKSL	840
DB	781	LNLYKSSRHLLAFINKFVQFIHKYTTYNAPAAISFLQKHADPLHDLSFSDNSDLVMLKSL	840
QY	841	LAGLSLPSRDDRDRGLDDEGEESAGSLPVSVSLFTPLTAEMAAPYMKRLRSRCQTYE	900
DB	841	LAGLSLPSRDDRDRGLDDEGEESAGSLPVSVSLFTPLTAEMAAPYMKRLRSRCQTYE	900
QY	901	DELEVSLDIDEMSRREPELISFSTMLQRLMSAECCCNLAFLSLAROMNSPSTAAAF	960
DB	901	DELEVSLDIDEMSRREPELISFSTMLQRLMSAECCCNLAFLSLAROMNSPSTAAAF	960
QY	961	LPTFMVCLGSDPEVYQTLRLNLPYALLCQEHAAVLLHRAFLVGYMGMDPSAQISEAL	1020
DB	961	LPTFMVCLGSDPEVYQTLRLNLPYALLCQEHAAVLLHRAFLVGYMGMDPSAQISEAL	1020
QY	1021	RIIHEAVM 1029	
DB	1021	RIIHEAVM 1029	

```

RESULT 7
US-09-929-769-7
; Sequence 7, Application US/09929769
; Publication No: US20030055224A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Mel-Olang
; APPLICANT: Polakis, Paul
; APPLICANT: Shou, Jianyong
; APPLICANT: Smith, Victoria
; APPLICANT: Soriano, Robert
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF TUMOR
; FILE REFERENCE: P5007R1-US
; CURRENT APPLICATION NUMBER: US/09/929, 769
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/089,653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/090,355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/104,257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/141,037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/20111
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: PCT/US00/00376
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PCT/US00/04342
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/13705
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/20118
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/888,257
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-929-769-7

Query Match      100.0%; Score 1029; DB 9; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 61 RMIRSEVLRLVDAALQDLPEQQLLEFVQSGFIPVSSSKILLQFLDQAVHPDQTLQONIM 120
DB 61 RMIRSEVLRLVDAALQDLPEQQLLEFVQSGFIPVSSSKILLQFLDQAVHPDQTLQONIM 120
QY 121 DKNYMAHLVEQVHERGASGQTFHSLLTASLPFRDSTEARPKPSSPEOPIGQRIHVG 180
DB 121 DKNYMAHLVEQVHERGASGQTFHSLLTASLPFRDSTEARPKPSSPEOPIGQRIHVG 180
QY 181 QLRVLGPEDDLQAGMFLQIFPLSPDRMSSPRVALALQALQGLARVQSGPEVPGI 240
DB 181 QLRVLGPEDDLQAGMFLQIFPLSPDRMSSPRVALALQALQGLARVQSGPEVPGI 240
QY 241 TVRYVQLATLILSSPHGALVMSHRSHPFLACPLLRQLQYQRCVPDGTGSSFLKVL 300
DB 241 TVRYVQLATLILSSPHGALVMSHRSHPFLACPLLRQLQYQRCVPDGTGSSFLKVL 300
QY 301 QMLQMLSPGVEGGLPRLAOLRLASQASAGRLSDVGGILLRLAEALAFRODLEVSSTV 360
DB 301 QMLQMLSPGVEGGLPRLAOLRLASQASAGRLSDVGGILLRLAEALAFRODLEVSSTV 360
QY 361 RAVIATLRSGEQCSVEPDLISKVLQGLIEVRSPLHEELTAFFSATADAASPPACKPVY 420
DB 361 RAVIATLRSGEQCSVEPDLISKVLQGLIEVRSPLHEELTAFFSATADAASPPACKPVY 420
QY 421 VVSSLILQEEEPFLAGKRGADGSLAVRIGPSSGGLVMDLEMDPEVSSCPDLQRL 480
DB 421 VVSSLILQEEEPFLAGKRGADGSLAVRIGPSSGGLVMDLEMDPEVSSCPDLQRL 480
QY 481 FSRKRGQAQVPSFRPYLLTLFTHOSWPTLHQCIVLGLGKSGEQFDPDSASLDFLMAC 540
DB 481 FSRKRGQAQVPSFRPYLLTLFTHOSWPTLHQCIVLGLGKSGEQFDPDSASLDFLMAC 540
QY 541 IHVPRIWQGRDQTPQKRREELVLRVGPBELLISVELILAEATRSODGCTAACSLIQAR 600
DB 541 IHVPRIWQGRDQTPQKRREELVLRVGPBELLISVELILAEATRSODGCTAACSLIQAR 600
QY 601 LPLILSCCGDDSVKRVTEHLSCGICQWGDVYGRRCROLLQLYQRELRVPEVL 660
DB 601 LPLILSCCGDDSVKRVTEHLSCGICQWGDVYGRRCROLLQLYQRELRVPEVL 660
QY 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRLAAVAHPDLLRL 720
DB 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRLAAVAHPDLLRL 720
QY 721 PMIAALLHGRTHLNFQFROONHLSCEFLAVGLLELLQPHVFRSEHGALMDCLSIRL 780
DB 721 PMIAALLHGRTHLNFQFROONHLSCEFLAVGLLELLQPHVFRSEHGALMDCLSIRL 780
QY 781 LNAVYKSSRLAALFINKFVOFIRKYYTYNAPAAISFLQKHADPLHDSFNSDLVMSKSL 840
DB 781 LNAVYKSSRLAALFINKFVOFIRKYYTYNAPAAISFLQKHADPLHDSFNSDLVMSKSL 840
QY 841 LAGLSLPSRDRDRLDGLDEGESESSAGSLPLVSVSLFTPLTAEMA PYMKRLSRGQTV 900
DB 841 LAGLSLPSRDRDRLDGLDEGESESSAGSLPLVSVSLFTPLTAEMA PYMKRLSRGQTV 900
QY 901 DLLEVLSDIDEMSRREILISFSTNLQRLMSSAEECCRNLAFLARSQNSPSTIAAF 960
DB 901 DLLEVLSDIDEMSRREILISFSTNLQRLMSSAEECCRNLAFLARSQNSPSTIAAF 960
QY 961 LPTFMVCLGSDQFEVYOTALRNLPEYALLQCEHAAVYLHRAFLVGMYGOMDPSAQISEAL 1020
DB 961 LPTFMVCLGSDQFEVYOTALRNLPEYALLQCEHAAVYLHRAFLVGMYGOMDPSAQISEAL 1020
QY 1021 RILHMEAVM 1029
DB 1021 RILHMEAVM 1029

```

```

QY 1 MHLVVHNAVYILTLGPRRADSEFOALDDIMPEEKPLTAFLVDTSEALILPMWKL 60
DB 1 MHLVVHNAVYILTLGPRRADSEFOALDDIMPEEKPLTAFLVDTSEALILPMWKL 60

```

```

RESULT 8
US-10-032-996-22
; Sequence 22, Application US/10032996
; Publication No. US20030054447A1

```

GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P293ORIC3
CURRENT APPLICATION NUMBER: US/10/032,996
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 22
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-10-032-996-22

Query Match 100.0% Score 1029, DB 9, Length 1029;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHLYVHAWYILLTTPRRADSEFOALDIWPEKPLPTALVOTSEALLPOMKL 60
DB 1 MHLYVHAWYILLTTPRRADSEFOALDIWPEKPLPTALVOTSEALLPOMKL 60
QY 61 RMIRSEVLRLVDAALODLEPOQLLFFVOSGIVVSSMSKLLQPLDVAHADPOTLCONIM 120
DB 61 RMIRSEVLRLVDAALODLEPOQLLFFVOSGIVVSSMSKLLQPLDVAHADPOTLCONIM 120
QY 121 DKNYMAHVEVOHERGASGGQTFHSLTASLPPRRDSTEAPKPKSSPEQIGGRIRVGT 180
DB 121 DKNYMAHVEVOHERGASGGQTFHSLTASLPPRRDSTEAPKPKSSPEQIGGRIRVGT 180
QY 181 QLRVLGPEDDLAGMFLQITPLSPDPKMOSSPPVALALQALGQELARVYGSSPEVPGI 240
DB 181 QLRVLGPEDDLAGMFLQITPLSPDPKMOSSPPVALALQALGQELARVYGSSPEVPGI 240
QY 241 TVRVLQALATLSSPHGALVMSHRSFLACPLRLQLOYORCVPODTGESSFLKVL 300
DB 241 TVRVLQALATLSSPHGALVMSHRSFLACPLRLQLOYORCVPODTGESSFLKVL 300
QY 301 OMLOWIDSPGVEGGLRAQLRMLASQASAGRRLSDVRCGLRLRAEALAEFRODEVYSTV 360
DB 301 OMLOWIDSPGVEGGLRAQLRMLASQASAGRRLSDVRCGLRLRAEALAEFRODEVYSTV 360
QY 361 RAVIATLRSGEOCSVEPDLISKVLOGLIEYRSHLEBELTAFTSATADAAPPACKPVY 420
DB 361 RAVIATLRSGEOCSVEPDLISKVLOGLIEYRSHLEBELTAFTSATADAAPPACKPVY 420
QY 421 VVSSLLQEEPLAGKPGADGSLFAVRIGPSSGLVIMLEMLDEPVVSSCDLQRL 480
DB 421 VVSSLLQEEPLAGKPGADGSLFAVRIGPSSGLVIMLEMLDEPVVSSCDLQRL 480
QY 481 FSRKRGGAQVPSFRPYLLTFTHOSWPTLHQICIRVLLGKREGRFEPASLDELTMAC 540
DB 481 FSRKRGGAQVPSFRPYLLTFTHOSWPTLHQICIRVLLGKREGRFEPASLDELTMAC 540
QY 541 IHVPRIMOGRODQTPQKRREELVLRVQGPPELLISVELILAEATRSDGDTAACSIIQAR 600
DB 541 IHVPRIMOGRODQTPQKRREELVLRVQGPPELLISVELILAEATRSDGDTAACSIIQAR 600
QY 601 LPILLSGCCGDDSVKRVTEHLSGCIQOMGDSVLRGRCDLLQVLRQRELVPVPEYL 660
DB 601 LPILLSGCCGDDSVKRVTEHLSGCIQOMGDSVLRGRCDLLQVLRQRELVPVPEYL 660
QY 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRLAVAHPLLLRL 720
DB 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACRLAVAHPLLLRL 720
QY 721 PMTALLHGRTHINPOEFROQNLSCFHVYLGLELLQHPVFSBHOAGAMDCLLSFIRL 780
DB 721 PMTALLHGRTHINPOEFROQNLSCFHVYLGLELLQHPVFSBHOAGAMDCLLSFIRL 780
QY 781 LUNYRKSSRLAFAINKFVOFIHKYTYNAPAAISFLQKHADPLHDSFNSDLVMLKSL 840
DB 781 LUNYRKSSRLAFAINKFVOFIHKYTYNAPAAISFLQKHADPLHDSFNSDLVMLKSL 840
QY 841 LAGLSLPSRDDRTRGLDEBGEESAGSLPLVSVSLFTPLTAEMAPYKRLSRQGYE 900
DB 841 LAGLSLPSRDDRTRGLDEBGEESAGSLPLVSVSLFTPLTAEMAPYKRLSRQGYE 900
QY 901 DLLEVLSDIDEMRRRPELISFSTNQLRMSAECCRLASLALRSQNSPSTAAAF 960
DB 901 DLLEVLSDIDEMRRRPELISFSTNQLRMSAECCRLASLALRSQNSPSTAAAF 960
QY 961 LPTFMVCLSGODEEVVQVOTALRNLPEVALCOEHAVALIHRAPLVGMYGQMDPSAQISEAL 1020
DB 961 LPTFMVCLSGODEEVVQVOTALRNLPEVALCOEHAVALIHRAPLVGMYGQMDPSAQISEAL 1020
QY 1021 RIILMEAVM 1029
DB 1021 RIILMEAVM 1029

Db 1021 RILHMEAVM 1029

RESULT 9

US-10-033-396-22
 ; Sequence 22, Application US/10033396
 ; Publication No. US20030077657A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnovers, Luc
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Collin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2930R1C4
 ; CURRENT APPLICATION NUMBER: US/10/033,396
 ; CURRENT FILING DATE: 2001-12-27
 ; PRIOR APPLICATION NUMBER: 60/095,325
 ; PRIOR FILING DATE: 1998-08-04
 ; PRIOR APPLICATION NUMBER: 60/112,851
 ; PRIOR FILING DATE: 1998-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,145
 ; PRIOR FILING DATE: 1998-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,511
 ; PRIOR FILING DATE: 1998-12-22
 ; PRIOR APPLICATION NUMBER: 60/115,558
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/115,565
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/115,733
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/119,341
 ; PRIOR FILING DATE: 1999-02-09
 ; PRIOR APPLICATION NUMBER: 60/119,537
 ; PRIOR FILING DATE: 1999-02-10
 ; PRIOR APPLICATION NUMBER: 60/119,965
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: 60/162,506
 ; PRIOR FILING DATE: 1999-10-29
 ; PRIOR APPLICATION NUMBER: 60/170,262
 ; PRIOR FILING DATE: 1999-12-09
 ; PRIOR APPLICATION NUMBER: 60/187,202
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: PCT/US99/12252
 ; PRIOR FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28634
 ; PRIOR FILING DATE: 1999-12-01
 ; PRIOR APPLICATION NUMBER: PCT/US99/28551
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/03565
 ; PRIOR FILING DATE: 2000-02-11
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: PCT/US00/05841
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/08439
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: PCT/US00/14941
 ; PRIOR FILING DATE: 2000-05-30
 ; PRIOR APPLICATION NUMBER: PCT/US00/15264
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/32678
 ; PRIOR FILING DATE: 2000-12-01
 ; NUMBER OF SEQ ID NOS: 38

; SEQ ID NO 22
 ; LENGTH: 1029
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-033-396-22
 Query Match 100.0%; Score 1029; DB 9; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MHILVHAMVILLTLGPRADSEFOALDITMPEKPLPTAFIVDSEELLLPDKTL	60
Db	1	MHILVHAMVILLTLGPRADSEFOALDITMPEKPLPTAFIVDSEELLLPDKTL	60
Qy	61	RMIRSEVRLVDALQDLEPQQLLFVQSGIPVSSMSKLLQFLDQAAVHPQTEQNIM	120
Db	61	RMIRSEVRLVDALQDLEPQQLLFVQSGIPVSSMSKLLQFLDQAAVHPQTEQNIM	120
Qy	121	DKNYMAHLYEVQHERGASGGOTFHSLLTASLPRRDSTEAPKPKSSPEQPIGQGRIRVGT	180
Db	121	DKNYMAHLYEVQHERGASGGOTFHSLLTASLPRRDSTEAPKPKSSPEQPIGQGRIRVGT	180
Qy	181	QLRVGPEDDLQAGMFLQIFPLSPDRMSSPPVALALQOALGQELARVVOGSEVPGI	240
Db	181	QLRVGPEDDLQAGMFLQIFPLSPDRMSSPPVALALQOALGQELARVVOGSEVPGI	240
Qy	241	TVRVQALATLTLSSPHGALVMSMRSHFLACPLLRQLCOYQRCVPODTGSSLEFLKVL	300
Db	241	TVRVQALATLTLSSPHGALVMSMRSHFLACPLLRQLCOYQRCVPODTGSSLEFLKVL	300
Qy	301	QMLQWLDSPGVGGPLRAQLRLASQASAGRLLSVRGGLRLAALAFRODLEVSSTV	360
Db	301	QMLQWLDSPGVGGPLRAQLRLASQASAGRLLSVRGGLRLAALAFRODLEVSSTV	360
Qy	361	RAVATLRSRGCSVEPPLISKVLQGLLEVRSPHLEELTFEFSADAAAPFPACKRVV	420
Db	361	RAVATLRSRGCSVEPPLISKVLQGLLEVRSPHLEELTFEFSADAAAPFPACKRVV	420
Qy	421	VVSSLLOEBEPLAGKRGADGSLAVALRGPSSGLLVDMLEMDPEVSSCPDQLRL	480
Db	421	VVSSLLOEBEPLAGKRGADGSLAVALRGPSSGLLVDMLEMDPEVSSCPDQLRL	480
Qy	481	FSRRKGKQAOVPSFRPLTLFTHQSSWPLTHOCIRVLGKSRQRDPASLDFLWAC	540
Db	481	FSRRKGKQAOVPSFRPLTLFTHQSSWPLTHOCIRVLGKSRQRDPASLDFLWAC	540
Qy	541	IHPRIWGRQRTPOKRRBELVLRVGPETLSVELLIAEFTSQGDTRACSLIOAR	600
Db	541	IHPRIWGRQRTPOKRRBELVLRVGPETLSVELLIAEFTSQGDTRACSLIOAR	600
Qy	601	LPLLSCCCGDESVKVTETHLSCIQQWGSVIGRRCDLLDLYLQRPFLRVVPEVL	660
Db	601	LPLLSCCCGDESVKVTETHLSCIQQWGSVIGRRCDLLDLYLQRPFLRVVPEVL	660
Qy	661	LHSEGAASSVCKLDGLHRTITLLADTSDSRALENBRADSMACRKAVALPILLHL	720
Db	661	LHSEGAASSVCKLDGLHRTITLLADTSDSRALENBRADSMACRKAVALPILLHL	720
Qy	721	PMIALLGRTHLNFOEPRQONHLSCLFVHGLLELLOPHYFRSHOGALWDCLSFTRL	780
Db	721	PMIALLGRTHLNFOEPRQONHLSCLFVHGLLELLOPHYFRSHOGALWDCLSFTRL	780
Qy	781	LLNTRKSSRHLLAFLNKEVQETIHKYTYNAPAAISFLQKADPPLHDSLFDSNDLVMLKSL	840
Db	781	LLNTRKSSRHLLAFLNKEVQETIHKYTYNAPAAISFLQKADPPLHDSLFDSNDLVMLKSL	840
Qy	841	LAGLSLPSRDRTDRLGDEBEESAGSLPLVYSGLFTPLTAAMAFYMKRLSGQIVE	900
Db	841	LAGLSLPSRDRTDRLGDEBEESAGSLPLVYSGLFTPLTAAMAFYMKRLSGQIVE	900
Qy	901	DLLEVLSDIDMSRRRPETLSFFSTNIDRLMSAEECCRNLAFLSIALMSNPSIAAF	960
Db	901	DLLEVLSDIDMSRRRPETLSFFSTNIDRLMSAEECCRNLAFLSIALMSNPSIAAF	960

QY 961 LPTFMKCSGDFEVVOTLRNLPEYALLCQEHAAVLHRAFLVGYGMDPSAQISEAL 1020
 DB 961 LPTFMKCSGDFEVVOTLRNLPEYALLCQEHAAVLHRAFLVGYGMDPSAQISEAL 1020
 QY 1021 RILHMEAVM 1029
 DB 1021 RILHMEAVM 1029
 RESULT 10
 US-10-033-246-22
 ; Sequence 22, Application US/10033246
 ; Patent No. US20020098505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2930R1C12
 ; CURRENT FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: 60/095,325
 ; PRIOR FILING DATE: 1998-08-04
 ; PRIOR APPLICATION NUMBER: 60/112,851
 ; PRIOR FILING DATE: 1998-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,145
 ; PRIOR FILING DATE: 1998-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,511
 ; PRIOR FILING DATE: 1998-12-22
 ; PRIOR APPLICATION NUMBER: 60/115,558
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/115,565
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/115,733
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/119,341
 ; PRIOR FILING DATE: 1999-02-09
 ; PRIOR APPLICATION NUMBER: 60/119,537
 ; PRIOR FILING DATE: 1999-02-10
 ; PRIOR APPLICATION NUMBER: 60/119,965
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: 60/162,506
 ; PRIOR FILING DATE: 1999-10-29
 ; PRIOR APPLICATION NUMBER: 60/170,262
 ; PRIOR FILING DATE: 1999-12-09
 ; PRIOR APPLICATION NUMBER: 60/187,202
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: PCT/US99/12252
 ; PRIOR FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28634
 ; PRIOR FILING DATE: 1999-12-01
 ; PRIOR APPLICATION NUMBER: PCT/US99/28551
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/03565
 ; PRIOR FILING DATE: 2000-02-11
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: PCT/US00/05841
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/08439
 ; PRIOR FILING DATE: 2000-03-30

QY PRIOR APPLICATION NUMBER: PCT/US00/14941
 ; PRIOR FILING DATE: 2000-05-30
 ; PRIOR APPLICATION NUMBER: PCT/US00/15264
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/32678
 ; PRIOR FILING DATE: 2000-12-01
 ; NUMBER OF SEQ ID NOS: 38
 ; SEQ ID NO 22
 ; LENGTH: 1029
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-033-246-22
 Query Match 100.0%; Score 1029; DB 12; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MHILVHAMVILTLTGPPRADSEFOALDIMEPEKPLPTAFVDTSEBALLPMLKL 60
 DB 1 MHILVHAMVILTLTGPPRADSEFOALDIMEPEKPLPTAFVDTSEBALLPMLKL 60
 QY 61 RMIRSEVLRYDAALDLEPQQLLFVQSGFIPVSSMSKLLQFLDQAVAHDPQTLEQNIM 120
 DB 61 RMIRSEVLRYDAALDLEPQQLLFVQSGFIPVSSMSKLLQFLDQAVAHDPQTLEQNIM 120
 QY 121 DKNYMAHVEVQHERGASGGQTFHSILTFASLPERRDSTEAPKPKSSPEOPICGRIRVGT 180
 DB 121 DKNYMAHVEVQHERGASGGQTFHSILTFASLPERRDSTEAPKPKSSPEOPICGRIRVGT 180
 QY 181 QLRVLGPEDDLACMFLQIFPLSPDPWMOSSPPRVALLQALGQELARYVQSSPEVPGI 240
 DB 181 QLRVLGPEDDLACMFLQIFPLSPDPWMOSSPPRVALLQALGQELARYVQSSPEVPGI 240
 QY 241 TVRVLQALATLLSSPHGALVMSMRSHPLACGLLQLOQYQYQVODPQCFSSLFKVL 300
 DB 241 TVRVLQALATLLSSPHGALVMSMRSHPLACGLLQLOQYQYQVODPQCFSSLFKVL 300
 QY 301 QMLQWLDSPGVEGGLRAQLRMLASQASGRRLSDVRGGLRLAEALAEAFRODELEVSSIV 360
 DB 301 QMLQWLDSPGVEGGLRAQLRMLASQASGRRLSDVRGGLRLAEALAEAFRODELEVSSIV 360
 QY 361 RAVIATLRSGEGCSEPDILSKVLQGLIEVRSFHEELLTAFFSAPADAAPPACKPVY 420
 DB 361 RAVIATLRSGEGCSEPDILSKVLQGLIEVRSFHEELLTAFFSAPADAAPPACKPVY 420
 QY 421 VVSSLLQEEEPPLAGKPGADGSLAVRGLGSSGLVWLEMLDPEVAVSSCPDLQRL 480
 DB 421 VVSSLLQEEEPPLAGKPGADGSLAVRGLGSSGLVWLEMLDPEVAVSSCPDLQRL 480
 QY 481 FSRKRGKGAQVPSFRPYLLTLTFTHQSWPTLHQCIRVLLGKSGREORFPASIDELWAC 540
 DB 481 FSRKRGKGAQVPSFRPYLLTLTFTHQSWPTLHQCIRVLLGKSGREORFPASIDELWAC 540
 QY 541 IHVPRIMOGDORTPOKRREELVRYQGPGLISLVELLAEATRSQDGTAACSLIQAR 600
 DB 541 IHVPRIMOGDORTPOKRREELVRYQGPGLISLVELLAEATRSQDGTAACSLIQAR 600
 QY 601 LPFLSCCGGDDSVKRVTEHLSGCIQOMGDSVLRGRCRDLQLQYLQRELVVPEYL 660
 DB 601 LPFLSCCGGDDSVKRVTEHLSGCIQOMGDSVLRGRCRDLQLQYLQRELVVPEYL 660
 QY 661 LHSGEAASSVCKLDGLIHREITLADTSDSRALENKADASMACRKLAVAHPLLRL 720
 DB 661 LHSGEAASSVCKLDGLIHREITLADTSDSRALENKADASMACRKLAVAHPLLRL 720
 QY 721 PMIAALLHGRTHLNFOEPRQONHLSGFHYLGLLELLQPRVFSSEHGALMDLSEFIRL 780
 DB 721 PMIAALLHGRTHLNFOEPRQONHLSGFHYLGLLELLQPRVFSSEHGALMDLSEFIRL 780
 QY 781 LUNYRKSSRHIAFINKFQFIHKYTTYNAPPAISFLQKHADPLHLSFDNSDLVMLKSL 840
 DB 781 LUNYRKSSRHIAFINKFQFIHKYTTYNAPPAISFLQKHADPLHLSFDNSDLVMLKSL 840

```

QY 841 LAGLSLPSRDDRDGLDEBGESESSAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTV 900
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
Db 841 LAGLSLPSRDDRDGRGLDEBGESESSAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTV 900
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
QY 901 DLEVLSDIDEMSRRELLISFSTNLQRLMSSAECCRNLAFLARSMONSSIAAF 960
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
Db 901 DLEVLSDIDEMSRRELLISFSTNLQRLMSSAECCRNLAFLARSMONSSIAAF 960
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
QY 961 LPTFMYCIGSDQFEVYOTALRNLPEYALLCOENAAVLHRAFLVGYGOMPSAQISEAL 1020
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
Db 961 LPTFMYCIGSDQFEVYOTALRNLPEYALLCOENAAVLHRAFLVGYGOMPSAQISEAL 1020
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 1021 RILHMEAVM 1029
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 1021 RILHMEAVM 1029
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678

```

```

; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 22
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-301-22
Query Match 100.0%; Score 1029; DB 12; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHILVHAMVILLTLGPPRADSEFOALLDIPFEKPLPAFLVDISEALLLPDWLKL 60
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
Db 1 MHILVHAMVILLTLGPPRADSEFOALLDIPFEKPLPAFLVDISEALLLPDWLKL 60
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
QY 61 RMIRSEVRLVDALQDLEPQALLLFVQSGFIPVSSMSKLLQFDQVAHHPQTEONIM 120
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
Db 61 RMIRSEVRLVDALQDLEPQALLLFVQSGFIPVSSMSKLLQFDQVAHHPQTEONIM 120
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
QY 121 DKNYMAHLVEYQHERGASGGQTFHSLTLASLPRRDSTEAPKPKSPPOPIGGRIRYGT 180
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
Db 121 DKNYMAHLVEYQHERGASGGQTFHSLTLASLPRRDSTEAPKPKSPPOPIGGRIRYGT 180
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 181 QLRVIGPDDDLAGMFLQIFLSPDRMSSSPRPALALQALQOLARVVOGSEVPGI 240
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 181 QLRVIGPDDDLAGMFLQIFLSPDRMSSSPRPALALQALQOLARVVOGSEVPGI 240
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 241 TVRYLQALATLTLSPHSGALVMSHSHFLACPLLRQICQYQRCVPODTGFSFLKYL 300
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 241 TVRYLQALATLTLSPHSGALVMSHSHFLACPLLRQICQYQRCVPODTGFSFLKYL 300
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 301 QMLQWLDSPGVGGPLRQRLMLASQASAGRRLLSDVRGGLRLALAFRODLEVYSTV 360
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 301 QMLQWLDSPGVGGPLRQRLMLASQASAGRRLLSDVRGGLRLALAFRODLEVYSTV 360
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 361 RAVATATLRSGGQSVPEPLISKVLQGLLEVRSPLHEELLTAFFSATADAASPPACKPVV 420
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 361 RAVATATLRSGGQSVPEPLISKVLQGLLEVRSPLHEELLTAFFSATADAASPPACKPVV 420
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 421 VVSSLQLEEPPLAGKKGADGSLAVRLGSSGLVDWLEMLDPEVSSCPDLQRL 480
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 421 VVSSLQLEEPPLAGKKGADGSLAVRLGSSGLVDWLEMLDPEVSSCPDLQRL 480
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 481 FSRKRGKQAOVPSFRPYLLTLFTHOSWPTLHOCIRYLKSKRQRPDSASDLFLMAC 540
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 481 FSRKRGKQAOVPSFRPYLLTLFTHOSWPTLHOCIRYLKSKRQRPDSASDLFLMAC 540
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 541 IHVPRIMGROORTPOKREELVLRVQGPPLISVELILAEFRSOGDPAACSLIOAR 600
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 541 IHVPRIMGROORTPOKREELVLRVQGPPLISVELILAEFRSOGDPAACSLIOAR 600
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 601 LPLILSCCGGDESVRYKTEHLISGCIQOWGDSVILGRRCRDLLDLYLQRPFLRVVPEVL 660
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 601 LPLILSCCGGDESVRYKTEHLISGCIQOWGDSVILGRRCRDLLDLYLQRPFLRVVPEVL 660
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 661 LHSEGAASSVCKLDGLIHRFTYLLADTSDSRALENRGADASMACKRLAVAHPLLLRL 720
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
Db 661 LHSEGAASSVCKLDGLIHRFTYLLADTSDSRALENRGADASMACKRLAVAHPLLLRL 720
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678
QY 721 PMIAALLHGRTHLNFQERQONHLSCLFHLVGLLELLQPHVFRSRHQALMDCLLSFRL 780
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/32678

```


Db 721 PMIAALLHGRTHLNFQEEFRQOHNLSCEFLVHLGLLELLOPHVFRSEHOGALMPCLLSIFRL 780
Qy 781 LLYAKSSSHLAAFINFKPQFTHKITYTNAAPASIFLOKHAPRLDSDNSDVLWKL 840
Db 781 LLYAKSSSHLAAFINFKPQFTHKITYTNAAPASIFLOKHAPRLDSDNSDVLWKL 840
Qy 841 LAGLSRSDRTDGLDEEGEESAGSLPLVSVSLFTPLAAEMAPMKRLSGQYVE 900
Db 841 LAGLSRSDRTDGLDEEGEESAGSLPLVSVSLFTPLAAEMAPMKRLSGQYVE 900
Qy 901 DLEVLSDIDEMSRRRPELISFFSTNLORLMSAECCRNLAFLALRSMQNSPIAAAF 960
Db 901 DLEVLSDIDEMSRRRPELISFFSTNLORLMSAECCRNLAFLALRSMQNSPIAAAF 960
Qy 961 LPTFNVCGSODEFVYVOTARLPEVALICQEHAAVLLHRAFLVGYMGMDPSAQISEAL 1020
Db 961 LPTFNVCGSODEFVYVOTARLPEVALICQEHAAVLLHRAFLVGYMGMDPSAQISEAL 1020
Qy 1021 RLHMEAVM 1029
Db 1021 RLHMEAVM 1029

RESULT 12
US-10-033-326-22
Sequence 22, Application US/1003326
Patent No. US20020098507A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1C8
CURRENT APPLICATION NUMBER: US/10/033,326
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03

PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 22
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-10-033-326-22
Query Match 100.0%; Score 1029; DB 12; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHILVYHAMVILLTTCPPRADDEFEQALDIMEPEKPPPTAFVDTSEALLPDMKL 60
Db 1 MHILVYHAMVILLTTCPPRADDEFEQALDIMEPEKPPPTAFVDTSEALLPDMKL 60
Qy 61 RMIRSEVRLVDAIADLEPPQOLLFLVQSFGRIPVSSMSKILQFLDQVAHADPOTLEQNTM 120
Db 61 RMIRSEVRLVDAIADLEPPQOLLFLVQSFGRIPVSSMSKILQFLDQVAHADPOTLEQNTM 120
Qy 121 DKNYMAHLVEVOHERGASGGQFHSLLTASLPRRSTEARPKSSPEQIGGRIRVGT 180
Db 121 DKNYMAHLVEVOHERGASGGQFHSLLTASLPRRSTEARPKSSPEQIGGRIRVGT 180
Qy 181 QLRVLGPEDDLAGMFLQIPIPLSPDPRMOSSPPRVALLQALGQELARVYQSSPEVPGI 240
Db 181 QLRVLGPEDDLAGMFLQIPIPLSPDPRMOSSPPRVALLQALGQELARVYQSSPEVPGI 240
Qy 241 TYRVLQALATLLSSPHGALVMSMRSHPFLACPLRLQLCQYQRCVQDGTGFSFLKVL 300
Db 241 TYRVLQALATLLSSPHGALVMSMRSHPFLACPLRLQLCQYQRCVQDGTGFSFLKVL 300
Qy 301 OMQLWDSPGVEGPRRLQRLMLASQASAGRRISDVGGGLRLAEALAFRODLEVYSSIV 360
Db 301 OMQLWDSPGVEGPRRLQRLMLASQASAGRRISDVGGGLRLAEALAFRODLEVYSSIV 360
Qy 361 RAVIATLRSGEQSVPEDLISKVLQGLIEVRSFHEBELLTAFPSATADAAPPACKPVY 420
Db 361 RAVIATLRSGEQSVPEDLISKVLQGLIEVRSFHEBELLTAFPSATADAAPPACKPVY 420
Qy 421 VVSSLLQEEBPLAGKPGADGSLAVRLGSSGLVDMLEMDPEVYSSCPDLOLRLL 480
Db 421 VVSSLLQEEBPLAGKPGADGSLAVRLGSSGLVDMLEMDPEVYSSCPDLOLRLL 480
Qy 481 FSRKRGGAQVPSFRPYLLTFTTHOSSWPTLHQCIRVLIGKRRORPDSASLDPLMC 540
Db 481 FSRKRGGAQVPSFRPYLLTFTTHOSSWPTLHQCIRVLIGKRRORPDSASLDPLMC 540
Qy 541 IHVPRIMOGRDQTPQKRREELVLRVQGPGLISLVELILAEATRSQDQDTAACSILQAR 600
Db 541 IHVPRIMOGRDQTPQKRREELVLRVQGPGLISLVELILAEATRSQDQDTAACSILQAR 600
Qy 601 LPLLLSCCGGDESVKRVTEHLSGCIOQMGDSVLRRCRDLLQLLYLRBELRVPEVY 660
Db 601 LPLLLSCCGGDESVKRVTEHLSGCIOQMGDSVLRRCRDLLQLLYLRBELRVPEVY 660

Db 601 LPLLSCCGDDSVRRKTEHLSCGICQWGDVLRRCRDLLQLYLQRELPVPEVL 660
Qy 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENGADASMACKRLAAVHPILLRHL 720
Db 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENGADASMACKRLAAVHPILLRHL 720
Qy 721 PMIAALLHGRTHLNFQEFROQNHLSCEFLHVLGELLELLQPHVFRSEHOGALMDCLLSFIRL 780
Db 721 PMIAALLHGRTHLNFQEFROQNHLSCEFLHVLGELLELLQPHVFRSEHOGALMDCLLSFIRL 780
Qy 781 LLMYRKSRRHLAIFINKFVQFIHKYITTYNAPAAISFLQKHADPLHDLSFQNSDLVMLKSL 840
Db 781 LLMYRKSRRHLAIFINKFVQFIHKYITTYNAPAAISFLQKHADPLHDLSFQNSDLVMLKSL 840
Qy 841 LAGSLSPSRDRTDRCGLDEGESESSAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTV 900
Db 841 LAGSLSPSRDRTDRCGLDEGESESSAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTV 900
Qy 901 DLLEVLSDIDEMSRREPEILSFSTNLQRLMSSAECCRNLAFLSRMNSPSIAAF 960
Db 901 DLLEVLSDIDEMSRREPEILSFSTNLQRLMSSAECCRNLAFLSRMNSPSIAAF 960
Qy 961 LPTFMVCLGSDQFEVYOTALRNLPEYALLCOEHAVALHRAFLVGMGMDPSAOISEAL 1020
Db 961 LPTFMVCLGSDQFEVYOTALRNLPEYALLCOEHAVALHRAFLVGMGMDPSAOISEAL 1020
Qy 1021 RILMEAVM 1029
Db 1021 RILMEAVM 1029

RESULT 13

US-09-984-271-235
; Sequence 235, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 235
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-271-235

Query Match 39.3%; Score 404; DB 9; Length 404;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 626 IQQMGDSVLGRRCRDLLQLYLQRELPVPEVLHSEGAASSVCKLDGLIHRFTLL 685
Db 1 IQQMGDSVLGRRCRDLLQLYLQRELPVPEVLHSEGAASSVCKLDGLIHRFTLL 60
Qy 686 ADSDSRALENGADASMACKRLAAVHPILLRHLPMIAALLHGRTHLNFQEFROQNHLS 745
Db 61 ADSDSRALENGADASMACKRLAAVHPILLRHLPMIAALLHGRTHLNFQEFROQNHLS 120
Qy 746 CFLHVLGELLELLQPHVFRSEHOGALMDCLLSFIRLLNMYKSSRHLLAIFINKFVQFIHKY 805
Db 746 CFLHVLGELLELLQPHVFRSEHOGALMDCLLSFIRLLNMYKSSRHLLAIFINKFVQFIHKY 805

Db 121 CFLHVLGELLELLQPHVFRSEHOGALMDCLLSFIRLLNMYKSSRHLLAIFINKFVQFIHKY 180
Qy 806 ITYNAPAAISFLQKHADPLHDLSFQNSDLVMLKSLAGSLPSRDRTDRCGLDEGESESS 865
Db 181 ITYNAPAAISFLQKHADPLHDLSFQNSDLVMLKSLAGSLPSRDRTDRCGLDEGESESS 240
Qy 866 SAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRREPEILSFST 925
Db 241 SAGSLPLVSVSLFTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRREPEILSFST 300
Qy 926 NLMRLMSSAECCRNLAFLSRMNSPSIAAFPLTFMVCIGSDQFEVYOTALRNLPE 985
Db 301 NLMRLMSSAECCRNLAFLSRMNSPSIAAFPLTFMVCIGSDQFEVYOTALRNLPE 360
Qy 986 YALLCOEHAVALHRAFLVGMGMDPSAOISEALRILHMEAVM 1029
Db 361 YALLCOEHAVALHRAFLVGMGMDPSAOISEALRILHMEAVM 404

RESULT 14

US-09-984-271-148
; Sequence 148, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 148
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (328)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-148

Query Match 31.8%; Score 327; DB 9; Length 328;
Best Local Similarity 100.0%; Pred. No. 1; 8e-286;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 703 MACRKLAAVHPILLRHLPMIAALLHGRTHLNFQEFROQNHLSCEFLHVLGELLELLQPHV 762
Db 1 MACRKLAAVHPILLRHLPMIAALLHGRTHLNFQEFROQNHLSCEFLHVLGELLELLQPHV 60
Qy 763 RSEHOGALMDCLLSFIRLLNMYKSSRHLLAIFINKFVQFIHKYITTYNAPAAISFLQKHAD 822
Db 61 RSEHOGALMDCLLSFIRLLNMYKSSRHLLAIFINKFVQFIHKYITTYNAPAAISFLQKHAD 120
Qy 823 PLHDLSFQNSDLVMLKSLAGSLPSRDRTDRCGLDEGESESSAGSLPLVSVSLFTPLT 882
Db 121 PLHDLSFQNSDLVMLKSLAGSLPSRDRTDRCGLDEGESESSAGSLPLVSVSLFTPLT 180
Qy 883 AAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRREPEILSFSTNLQRLMSSAECCRNLA 942
Db 181 AAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRREPEILSFSTNLQRLMSSAECCRNLA 240
Qy 943 FSLALRMONSPSIAAFPLTFMVCIGSDQFEVYOTALRNLPEYALLCOEHAVALHRAFL 1002
Db 943 FSLALRMONSPSIAAFPLTFMVCIGSDQFEVYOTALRNLPEYALLCOEHAVALHRAFL 1002

Db 241 FSLAIRSMQNSPSIAAFLPTMYCLGSODEFVOTALNLPAYALLCOEHAVALLHRAF 300
Qy 1003 LVGMYGOMDPSAQISEALRIHMEAVM 1029
|||||
Db 301 LVGMYGOMDPSAQISEALRIHMEAVM 327

Job time : 34 secs

RESULT 15
US-09-764-891-4155
; Sequence 4155, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4155
; LENGTH: 161
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (40)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (94)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (98)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (139)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (141)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4155

Query Match 4.2% Score 43; DB 9; Length 161;
Best Local Similarity 100.0%; Pred. No. 6.1e-31;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Qy 851 DRTDRGLDEGEERSSAGSLPLVSVSLFTPLTAEMAPYMKRL 893
|||||
Db 41 DRTDRGLDEGEERSSAGSLPLVSVSLFTPLTAEMAPYMKRL 83

Search completed: May 7, 2003, 19:47:23

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:42:59 ; Search time 48 Seconds
(without alignments)
4417.135 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 1029
Sequence: 1 MHILVHNAVYILTLGPRA.....MDPSAQLSEALRIHMEAYM 1029

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database:

SPTREMBL.21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	928	90.2	1377	4 Q9P2A8	Q9P2A8 homo sapien
2	809	78.6	812	4 Q96D36	Q96D36 homo sapien
3	698	67.8	698	4 Q8WV40	Q8WV40 homo sapien
4	633	61.5	640	4 Q9Y3W8	Q9Y3W8 homo sapien
5	260	25.3	673	4 Q9BT91	Q9BT91 homo sapien
6	100	9.7	310	4 Q9NTD1	Q9NTD1 homo sapien
7	36	3.5	818	11 Q91Z01	Q91Z01 mus musculu
8	18	1.7	162	11 Q9CTF7	Q9CTF7 mus musculu
9	9	0.9	241	16 Q9ZKNI	Q9ZKNI delinococcus
10	9	0.9	327	16 Q91345	Q91345 pseudomonas
11	8	0.8	81	3 Q9Y7G3	Q9Y7G3 coprinus co
12	8	0.8	151	4 Q9UG55	Q9UG55 homo sapien
13	8	0.8	154	4 Q9EPV0	Q9EPV0 azoarcus ey
14	8	0.8	157	5 Q9XTY5	Q9XTY5 caenorhabdi
15	8	0.8	207	5 Q9VRE4	Q9VRE4 dirosophila
16	8	0.8	226	2 Q8VWZ8	Q8VWZ8 rhodobacter

17	8	0.8	249	2 Q8VR19	Q8VR19 myxococcus
18	8	0.8	253	11 Q54831	Q54831 mus musculu
19	8	0.8	267	11 Q35333	Q35333 mus musculu
20	8	0.8	272	17 Q97ZT6	Q97ZT6 sulfolobus
21	8	0.8	279	16 Q37R19	Q37R19 streptococ
22	8	0.8	311	5 Q9W4H0	Q9W4H0 dirosophila
23	8	0.8	320	16 Q8XXM4	Q8XXM4 raietonia s
24	8	0.8	345	5 Q44916	Q44916 caenorhabdi
25	8	0.8	379	17 Q8ZVJ6	Q8ZVJ6 pyrobaculum
26	8	0.8	386	16 Q99WU9	Q99WU9 staphylococ
27	8	0.8	423	16 Q33313	Q33313 mycobacteri
28	8	0.8	436	5 Q95T23	Q95T23 dirosophila
29	8	0.8	490	10 Q94CS4	Q94CS4 oryza sativ
30	8	0.8	519	2 Q8RP20	Q8RP20 rhodovulm
31	8	0.8	543	2 Q8RMJ1	Q8RMJ1 leptospira
32	8	0.8	552	4 Q8WU59	Q8WU59 homo sapien
33	8	0.8	593	16 Q8Z1W5	Q8Z1W5 yersinia pe
34	8	0.8	640	16 Q9RX24	Q9RX24 delinococcus
35	8	0.8	672	6 Q9BE88	Q9BE88 macaca fasc
36	8	0.8	677	11 Q8R251	Q8R251 mus musculu
37	8	0.8	705	4 Q9H9S2	Q9H9S2 homo sapien
38	8	0.8	728	2 Q9RH22	Q9RH22 allicyclobac
39	8	0.8	780	13 Q9YGM3	Q9YGM3 figu rubrip
40	8	0.8	780	13 Q9PWQ1	Q9PWQ1 figu rubrip
41	8	0.8	789	5 Q9H8F2	Q9H8F2 trypanosoma
42	8	0.8	853	4 Q9H8M5	Q9H8M5 homo sapien
43	8	0.8	882	4 Q9BOM8	Q9BOM8 homo sapien
44	8	0.8	893	4 Q9NDM1	Q9NDM1 homo sapien
45	8	0.8	913	17 Q8TVS4	Q8TVS4 methanopyru

ALIGNMENTS

RESULT 1

Q9P2A8 PRELIMINARY; PRT; 1377 AA.
AC Q9P2A8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE KIAA1440 protein (Fragment).
GN KIAA1440.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.,
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL; AB037861; BAA92678.1;
FT NON-TER
SQ SEQUENCE 1377 AA; 153070 MW; CE70589EF594FE7E CRC64;

Query Match 90.2%; Score 928; DB 4; Length 1377;
Best local similarity 99.9%; Pred. No. 0;
Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHILVHNAVYILTLGPRAADSEFQALDIWPEKRPPTAFVLTSEALLPWLKL 60
DB 349 MHILVHNAVYILTLGPRAADSEFQALDIWPEKRPPTAFVLTSEALLPWLKL 408
QY 61 RMIRSEVLRVDAALDDLEPQQLLFVQSGFIVSSMSKLPDQAVAHDPOTLEONIM 120
DB 409 RMIRSEVLRVDAALDDLEPQQLLFVQSGFIVSSMSKLPDQAVAHDPOTLEONIM 468
QY 121 DKNYMAHVEVQHERGASGGOTFHSILTLASLPERRDSTAPKPKSSPEPIGGRIRVGT 180

Db 469 DKNYMALVYQHGRSGSGQTFTSLTASLPKRDSTPAKPKPSPEQPIGQRIIVGT 528
 QY 181 QLRVLPEDDLAGMFLQIFLSPDPKMOSSPPRYVALALQOALGOELARVVOGPEVGI 240
 Db 529 QLRVLPEDDLAGMFLQIFLSPDPKMOSSPPRYVALALQOALGOELARVVOGPEVGI 588
 QY 241 TVRYLQALATLSSPHGALVMSMHRSHFLACPLLRQLCYORCPDGTGSSFLKVL 300
 Db 589 TVRYLQALATLSSPHGALVMSMHRSHFLACPLLRQLCYORCPDGTGSSFLKVL 648
 QY 301 QMLQWLDSPVEGGPLRAOLRLMASQASAGRRLSDVVGGLRLAEALAFQDLEVSSTV 360
 Db 649 QMLQWLDSPVEGGPLRAOLRLMASQASAGRRLSDVVGGLRLAEALAFQDLEVSSTV 708
 QY 361 RAVIATLRSGEQCSVEPDLISKVQLGIEVRSPLHEELLAFATATDAASPACRPV 420
 Db 709 RAVIATLRSGEQCSVEPDLISKVQLGIEVRSPLHEELLAFATATDAASPACRPV 768
 QY 421 VVSSLQEEEPPLAGKPGADGSLAVRLGPSSGLLVDMLEMLDPEVSSCPDLQLRL 480
 Db 769 VVSSLQEEEPPLAGKPGADGSLAVRLGPSSGLLVDMLEMLDPEVSSCPDLQLRL 828
 QY 481 FSRKKGQAOVSPFRYLLTLFTHOSWPTLHOCIRVLLGKSPREOFDSASIDELMAC 540
 Db 829 FSRKKGQAOVSPFRYLLTLFTHOSWPTLHOCIRVLLGKSPREOFDSASIDELMAC 888
 QY 541 IHVPRIWQGRDQTPOKRREELVLRVGPPELISVELILAEATRSODGDTAACSLIQAR 600
 Db 889 IHVPRIWQGRDQTPOKRREELVLRVGPPELISVELILAEATRSODGDTAACSLIQAR 948
 QY 601 LPLLLSCCGDDSVKRVYTHLSGCIQOMGDSVYGRRCRDLLOLYQRELRVPEVL 660
 Db 949 LPLLLSCCGDDSVKRVYTHLSGCIQOMGDSVYGRRCRDLLOLYQRELRVPEVL 1008
 QY 661 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACKRLAVALPILLRL 720
 Db 1009 LHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGADASMACKRLAVALPILLRL 1068
 QY 721 PMIAALLHGRTHLNFQFROONHLSCLHVLGLLELLQPHVFRSEHOGALMDCILSFIRL 780
 Db 1069 PMIAALLHGRTHLNFQFROONHLSCLHVLGLLELLQPHVFRSEHOGALMDCILSFIRL 1128
 QY 781 LMYRKRSSRLAFAINKFVQFIKYYTYNAPAAISFLQKADPHLSDSPNSDLVMLKSL 840
 Db 1129 LMYRKRSSRLAFAINKFVQFIKYYTYNAPAAISFLQKADPHLSDSPNSDLVMLKSL 1188
 QY 841 LAGLSLPSRDDRTDRGLDEGESEESSAGSLPLVSVSLFTPLTAEMAPYMKRSLRGOTVE 900
 Db 1189 LAGLSLPSRDDRTDRGLDEGESEESSAGSLPLVSVSLFTPLTAEMAPYMKRSLRGOTVE 1248
 QY 901 DLLEVLSDIDEMSRRPETLSFSTNLQRLMSSAECCRNILAFSLARSMQNSPSIAAAF 960
 Db 1249 DLLEVLSDIDEMSRRPETLSFSTNLQRLMSSAECCRNILAFSLARSMQNSPSIAAAF 1308
 QY 961 LPTFMVCLGSDQDEEVQVOTALRNLPEYALLQOEHAVALLHRAFLVGMGQMDPSAQISEAL 1020
 Db 1309 LPTFMVCLGSDQDEEVQVOTALRNLPEYALLQOEHAVALLHRAFLVGMGQMDPSAQISEAL 1368
 QY 1021 RILHMEAVM 1029
 Db 1369 RILHMEAVM 1377

RESULT 2

ID 096D36 PRELIMINARY: PRT: 812 AA.
 AC 096D36:
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DE Hypothetical 90.0 kDa protein (Fragment).
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LYMPH;
 RA Strausberg R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 KW EMBL: BC013367; AAH13367.1;
 DR Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 812 AA: 89996 MW: 5054AE099C15BFB4 CRC64;

Query Match 78.6%; Score 809; DB 4; Length 812;

Best local similarity 100.0%; Pred. No. 0;

Matches 809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 221 QALGOELARVVOGPEVPGITVRVQLATLTLSSPHGALVMSMHRSHFLACPLLRQLCQ 280
 Db 4 QALGOELARVVOGPEVPGITVRVQLATLTLSSPHGALVMSMHRSHFLACPLLRQLCQ 63
 QY 281 YORCPDGTGSSFLKVLQMLQWLDSPVEGGPLRAOLRLMASQASAGRRLSDVVGGL 340
 Db 64 YORCPDGTGSSFLKVLQMLQWLDSPVEGGPLRAOLRLMASQASAGRRLSDVVGGL 123
 QY 341 LRLAEALAFQDLEVSSTVRAVATLRSGEQCSVEPDLISKVQLGIEVRSPLHEELLT 400
 Db 124 LRLAEALAFQDLEVSSTVRAVATLRSGEQCSVEPDLISKVQLGIEVRSPLHEELLT 183
 QY 401 AFPSATDAASPACRPVAVVSSLLQEEEPPLAGKPGADGSLAVRLGPSSGLLVDM 460
 Db 184 AFPSATDAASPACRPVAVVSSLLQEEEPPLAGKPGADGSLAVRLGPSSGLLVDM 243
 QY 461 LEMLDPEVSSCPDLQLRLFLSRKKGQAOVSPFRYLLTLFTHOSWPTLHOCIRVLL 520
 Db 244 LEMLDPEVSSCPDLQLRLFLSRKKGQAOVSPFRYLLTLFTHOSWPTLHOCIRVLL 303
 QY 521 GKSREQFDPDSASIDELMACIHVPRIWQGRDQTPOKRREELVLRVGPPELISVELILA 580
 Db 304 GKSREQFDPDSASIDELMACIHVPRIWQGRDQTPOKRREELVLRVGPPELISVELILA 363
 QY 581 EAETRSODGDTAACSLIQARLPLLLSCCGDDSVKRVYTHLSGCIQOMGDSVYGRRCRD 640
 Db 364 EAETRSODGDTAACSLIQARLPLLLSCCGDDSVKRVYTHLSGCIQOMGDSVYGRRCRD 423
 QY 641 LLLQLYQRELRVPEVLLHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGAD 700
 Db 424 LLLQLYQRELRVPEVLLHSEGAASSVCKLDGLIHRFTLLADTSDSRALENRGAD 483
 QY 701 ASMACRKLAAVHPILLRLPMIAALLHGRTHLNFQFROONHLSCLHVLGLLELLQPH 760
 Db 484 ASMACRKLAAVHPILLRLPMIAALLHGRTHLNFQFROONHLSCLHVLGLLELLQPH 543
 QY 761 VFRSEHOGALMDCILSFIRLLNRYKSSRLAFAINKFVQFIKYYTYNAPAAISFLQKH 820
 Db 544 VFRSEHOGALMDCILSFIRLLNRYKSSRLAFAINKFVQFIKYYTYNAPAAISFLQKH 603
 QY 821 ADPLHDSFQNSDLVMLKSLLAGLSLPSRDDRTDRGLDEGESEESSAGSLPLVSVSLFT 880
 Db 604 ADPLHDSFQNSDLVMLKSLLAGLSLPSRDDRTDRGLDEGESEESSAGSLPLVSVSLFT 663
 QY 881 LTAEMAPYMKRSLRGOTVEDLLEVLSDIDEMSRRPETLSFSTNLQRLMSSAECCRN 940
 Db 664 LTAEMAPYMKRSLRGOTVEDLLEVLSDIDEMSRRPETLSFSTNLQRLMSSAECCRN 723
 QY 941 LAFSLARSMQNSPSIAAALPTFMVCLGSDQDEEVQVOTALRNLPEYALLQOEHAVALLHR 1000
 Db 724 LAFSLARSMQNSPSIAAALPTFMVCLGSDQDEEVQVOTALRNLPEYALLQOEHAVALLHR 783
 QY 1001 AFLVGMGQMDPSAQISEALRILHMEAVM 1029
 Db 784 AFLVGMGQMDPSAQISEALRILHMEAVM 812

RESULT 3
 08WV40 PRELIMINARY: PRT: 698 AA.
 AC 08WV40:
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical 77.7 kDa protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKIN;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC018777; AAH18777.1;
 KW Hypothetical protein.
 FT NON_TER
 SO SEQUENCE 698 AA; 77709 MW; C917828091BF4778 CRC64;

Query Match 67.8%; Score 698; DB 4; Length 698;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 698; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 RLSDVGGGLRLAEALAFRODLEVVSTVRVATLRLSGECCSVPEPLISKVLQGLIEVR 391
 DB 1 RLSDVGGGLRLAEALAFRODLEVVSTVRVATLRLSGECCSVPEPLISKVLQGLIEVR 60
 QY 392 SPHEELLTAFFSATADASPFPACKPVVVSSLLQEEPLAGKPGADGSLFAVRLG 451
 DB 61 SPHEELLTAFFSATADASPFPACKPVVVSSLLQEEPLAGKPGADGSLFAVRLG 120
 QY 452 PSSGLVDMLEMDPEVVSSCPDQLRLIFSRKKGGAQVPSFRPYLLTLFTHOSSWPT 511
 DB 121 PSSGLVDMLEMDPEVVSSCPDQLRLIFSRKKGGAQVPSFRPYLLTLFTHOSSWPT 180
 QY 512 LHOCIRVLLGSRQRPDSASIDFLMACIHVPRIMOGRODTPQKRREELVLRQGPRL 571
 DB 181 LHOCIRVLLGSRQRPDSASIDFLMACIHVPRIMOGRODTPQKRREELVLRQGPRL 240
 QY 572 ISIVELIAEAETRSQDQDFAACSLIQARLPILLSGCCGDESVRKYTEHLSGCIQOMGD 631
 DB 241 ISIVELIAEAETRSQDQDFAACSLIQARLPILLSGCCGDESVRKYTEHLSGCIQOMGD 300
 QY 632 SVLGRCDLLQLQYLRPELRVVPVYLLHSEGAASSVCKLDGLIHRFTLLADTSDS 691
 DB 301 SVLGRCDLLQLQYLRPELRVVPVYLLHSEGAASSVCKLDGLIHRFTLLADTSDS 360
 QY 692 RALERGADSMACRKLAVAHPLLLRLPMTAALLHGRTHLNPOEFOQNHLSCFLHVL 751
 DB 361 RALERGADSMACRKLAVAHPLLLRLPMTAALLHGRTHLNPOEFOQNHLSCFLHVL 420
 QY 752 GULELLQPHVFRSEHOGALMDCILSFIRLLNLYRKSSRHIAAFINKFVOFIHKYITYNAP 811
 DB 421 GULELLQPHVFRSEHOGALMDCILSFIRLLNLYRKSSRHIAAFINKFVOFIHKYITYNAP 480
 QY 812 LAISFLQKADPLHDLSFDNSDLVMLKSLAGLSLPSRDRTDRLDEGEESAGSLP 871
 DB 481 LAISFLQKADPLHDLSFDNSDLVMLKSLAGLSLPSRDRTDRLDEGEESAGSLP 540
 QY 872 LVSVSLFPLTAEMAPYMKRLSGQTVEDLLEVLSDIDEMSRRPPELISFSTINLQIRM 931
 DB 541 LVSVSLFPLTAEMAPYMKRLSGQTVEDLLEVLSDIDEMSRRPPELISFSTINLQIRM 600
 QY 932 SSAECCNLAFFSLALRSMONSPSIAAFLPTFMYCLGSODEFVVQATALRNLPETALLCO 991
 DB 601 SSAECCNLAFFSLALRSMONSPSIAAFLPTFMYCLGSODEFVVQATALRNLPETALLCO 660
 QY 992 EHAVALLHRAFLVGMYGOMDPSAOISEALRLIHEAVM 1029
 DB 660 EHAVALLHRAFLVGMYGOMDPSAOISEALRLIHEAVM 1029

DB 661 EHAVALLHRAFLVGMYGOMDPSAOISEALRLIHEAVM 698

RESULT 4
 09Y3W8 PRELIMINARY: PRT: 640 AA.
 AC 09Y3W8:
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 71.4 kDa protein (Fragment).
 OS Homo sapiens (Human).
 GN DKF2PS86J0619.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UTERUS;
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL050110; CAB43278.1;
 KW Hypothetical protein.
 FT NON_TER
 SO SEQUENCE 640 AA; 71375 MW; 1AC87FC2075B21A CRC64;

Query Match 61.5%; Score 633; DB 4; Length 640;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 397 ELTTAFESATADASPFPACKPVVVSSLLQEEPLAGKPGADGSLFAVRLGSSGL 456
 DB 8 ELTTAFESATADASPFPACKPVVVSSLLQEEPLAGKPGADGSLFAVRLGSSGL 67
 QY 457 LVDWLEMDPEVVSSCPDQLRLIFSRKKGGAQVPSFRPYLLTLFTHOSSWPTLHOCI 516
 DB 68 LVDWLEMDPEVVSSCPDQLRLIFSRKKGGAQVPSFRPYLLTLFTHOSSWPTLHOCI 127
 QY 517 RVLILKSRQRPDSASIDFLMACIHVPRIMOGRODTPQKRREELVLRQGPRL 576
 DB 128 RVLILKSRQRPDSASIDFLMACIHVPRIMOGRODTPQKRREELVLRQGPRL 187
 QY 577 LTLAEETRSQDQDFAACSLIQARLPILLSGCCGDESVRKYTEHLSGCIQOMGDSVLGR 636
 DB 188 LTLAEETRSQDQDFAACSLIQARLPILLSGCCGDESVRKYTEHLSGCIQOMGDSVLGR 247
 QY 637 RCRDILLQYLRPELRVVPVYLLHSEGAASSVCKLDGLIHRFTLLADTSDSRALEN 696
 DB 248 RCRDILLQYLRPELRVVPVYLLHSEGAASSVCKLDGLIHRFTLLADTSDSRALEN 307
 QY 697 RGADSMACRKLAVAHPLLLRLPMTAALLHGRTHLNPOEFOQNHLSCFLHVL 756
 DB 308 RGADSMACRKLAVAHPLLLRLPMTAALLHGRTHLNPOEFOQNHLSCFLHVL 420
 QY 757 LQPHVFRSEHOGALMDCILSFIRLLNLYRKSSRHIAAFINKFVOFIHKYITYNAP 816
 DB 427 LQPHVFRSEHOGALMDCILSFIRLLNLYRKSSRHIAAFINKFVOFIHKYITYNAP 480
 QY 817 LQKADPLHDLSFDNSDLVMLKSLAGLSLPSRDRTDRLDEGEESAGSLP 876
 DB 487 LQKADPLHDLSFDNSDLVMLKSLAGLSLPSRDRTDRLDEGEESAGSLP 540
 QY 877 LFTPLTAEMAPYMKRLSGQTVEDLLEVLSDIDEMSRRPPELISFSTINLQIRM 936
 DB 547 LFTPLTAEMAPYMKRLSGQTVEDLLEVLSDIDEMSRRPPELISFSTINLQIRM 600
 QY 937 CCRNLAFSLALRSMONSPSIAAFLPTFMYCLGSODEFVVQATALRNLPETALLCOEHAAY 996
 DB 607 CCRNLAFSLALRSMONSPSIAAFLPTFMYCLGSODEFVVQATALRNLPETALLCOEHAAY 660
 QY 997 LTHRAFLVGMYGOMDPSAOISEALRLIHEAVM 1029
 DB 608 LTHRAFLVGMYGOMDPSAOISEALRLIHEAVM 1029

```

RESULT 5
O9BT91
ID 09BT91 PRELIMINARY; PRT; 673 AA.
AC 09BT91:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE Hypothetical 75.1 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004286; AA04286.1; -
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 673 AA; 75055 MW; 6CE17654F5BA3DE CRC64;

Query Match
Best Local Similarity 25.3%; Score 260; DB 4; Length 673;
Matches 360; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MHLVYHAWVILTLGPPRADSEFOALDIMEPEEKPLPTAFVDTSEALLPWLKL 60
DB 304 MHLVYHAWVILTLGPPRADSEFOALDIMEPEEKPLPTAFVDTSEALLPWLKL 363
OY 61 RMIRSEVRLVDAALDLEPQQLLEVOFGIPVSSSKLLQFLDQAVADHPOTLEONIM 120
DB 364 RMIRSEVRLVDAALDLEPQQLLEVOFGIPVSSSKLLQFLDQAVADHPOTLEONIM 423
OY 121 DKYVMAHVLVEVHERGSGGOTFHSULTASLPPRODTEAPKPKSSPEQPIGGRIRVGT 180
DB 424 DKYVMAHVLVEVHERGSGGOTFHSULTASLPPRODTEAPKPKSSPEQPIGGRIRVGT 483
OY 181 QLRVLGPEDDLAGMFLQIFPLSPDPKMOSSPPRVALALQALGQELARVQSGPEVPGI 240
DB 484 QLRVLGPEDDLAGMFLQIFPLSPDPKMOSSPPRVALALQALGQELARVQSGPEVPGI 543
OY 241 TVAVYLAALILSSPHGALVMSHSHFLACPLRLQCYQRCVPODTGFSFLKVL 300
DB 544 TVAVYLAALILSSPHGALVMSHSHFLACPLRLQCYQRCVPODTGFSFLKVL 603
OY 301 QMLQWLDSPVEGGPLRAQLRMLASQASAGRRLSDVAGGLRLAEALAFRODLEVVSSTV 360
DB 604 QMLQWLDSPVEGGPLRAQLRMLASQASAGRRLSDVAGGLRLAEALAFRODLEVVSSTV 663
OY 361 R 361
DB 664 R 664

```

```

RA Ansoorge W., Wirkner U., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137358; CAB70710.1; -
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 310 AA; 34728 MW; EDBA23F1DOC08521 CRC64;

Query Match
Best Local Similarity 9.7%; Score 100; DB 4; Length 310;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 844 LSLPSRDRDRGIDGEGEESNAGSLPLVSVSLFTPLTAABAPYKRLRSRGQYEDLL 903
DB 89 LSLPSRDRDRGIDGEGEESNAGSLPLVSVSLFTPLTAABAPYKRLRSRGQYEDLL 148
OY 904 EVLSDIDEMSRRPPELISFSTNLRMSAECCRNIAF 943
DB 149 EVLSDIDEMSRRPPELISFSTNLRMSAECCRNIAF 168

RESULT 7
O91201
ID 091201 PRELIMINARY; PRT; 818 AA.
AC 091201:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE Hypothetical 90.9 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010333; AAH10333.1; -
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 818 AA; 90949 MW; A5128C8692BB7663 CRC64;

Query Match
Best Local Similarity 3.5%; Score 36; DB 11; Length 818;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 529 DPSASIDFLWACIHVPRWGRDQRTPOKRREELVL 564
DB 318 DPSASIDFLWACIHVPRWGRDQRTPOKRREELVL 353

RESULT 8
O9CTE7
ID 09CTE7 PRELIMINARY; PRT; 162 AA.
AC 09CTE7:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE 1110015K06R1k protein (Fragment).
GN 1110015K06R1K
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shidata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glassi C., King B., Kochiwa H.,

```


RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirni L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection.",
 RL Nature 409:685-690(2001).
 DR EMBL: AK003728; BAB22963.1; -
 DR MGD: MGI:1915760; 1110015K06R1k.
 FT NON_TER 1
 SO SEQUENCE 162 AA; 17939 MW; E343A7CF040CEC7A CRC64;

Query Match 1.7%; Score 18; DB 11; Length 162;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 988 LICEHAALVLRAPLVG 1005
 |||||
 DB 121 LICEHAALVLRAPLVG 138

RESULT 9

Q9RZNI PRELIMINARY; PRT; 241 AA.
 AC Q9RZNI;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE DNA-binding stress response protein, DPS family.
 GN DRB0092.
 OS Deinococcus radiodurans.
 OC Plasmid M1.
 CC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 CC Deinococcaceae; Deinococcus.
 CX NCBI_TaxID=1299;
 OX 11
 RN SEQUENCE FROM N.A.
 RP STRAIN=R1.
 RC MEDLINE=20036896; PubMed=10567266;
 RX White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Motif K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vanatvean J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.,
 RA "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1.",
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001826; AAF12541.1; -
 DR TIGR: DRB0092; -
 DR InterPro: IPR002177; DPS.
 DR Pfam: PF02047; DPS; 1.
 DR Prodom: PD149803; DPS; 1.
 KW DNA-binding; Plasmid; Complete proteome.
 SO SEQUENCE 241 AA; 26089 MW; 18F6C6F5D629E75 CRC64;

Query Match 0.9%; Score 9; DB 16; Length 241;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 419 VVVVSSLL 427
 |||||
 DB 8 VVVVSSLL 16

RESULT 10
 O91345 PRELIMINARY; PRT; 327 AA.
 AC O91345;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hypothetical protein PA1680.
 GN PA1680.
 OS Pseudomonas aeruginosa.
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 OX NCBI_TaxID=287;
 OX 11
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.",
 RL Nature 406:959-964(2000).
 DR EMBL: AE004595; AAC05069.1; -
 DR InterPro: IPR000734; Lipase.
 DR InterPro: IPR000379; Ser_estrs_site.
 DR PROSITE: PS00120; LIPASE_SER; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 327 AA; 35210 MW; C3DD0BA9351A1E67 CRC64;

Query Match 0.9%; Score 9; DB 16; Length 327;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 LRLRLAALA 348
 |||||
 DB 81 LRLRLAALA 89

RESULT 11
 O9Y7G3 PRELIMINARY; PRT; 81 AA.
 AC O9Y7G3;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Cop c1 allergen (Fragment).
 GN COP C1.
 OS Coprinus comatus (Shaggy mane).
 CC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 CC Agaricales; Agaricaceae; Coprinus.
 OX NCBI_TaxID=56187;
 OX 11
 RN SEQUENCE FROM N.A.
 RP TISSUE=CAP;
 RC MEDLINE=99414166; PubMed=10482839;
 RX Brander K.A., Bobely P., Cramer R., Pichler W.J., Heibling A.,
 RA "IGE-binding proliferative responses and skin test reactivity to Cop c
 RT 1, the first recombinant allergen from the basidiomycete Coprinus
 comatus.",
 RL J. Allergy Clin. Immunol. 104:630-636(1999).
 DR EMBL: AJ132235; CAB39376.1; -
 DR NON_TER 1
 SO SEQUENCE 81 AA; 8967 MW; E9CD4842DA24BB70 CRC64;

Query Match 0.8%; Score 8; DB 3; Length 81;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 710 VAHPLELL 717

Db 35 VARPLRL 42

RESULT 12

09UG55 PRELIMINARY; PRT; 151 AA.
AC 09UG55;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hypochemical 16.9 kDa protein (Fragment).
GN DKFZP34A091.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RA Poustika, A., Klein M., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL050372; CAB62530.1;
DR InterPro: IPR000120; Amidase.
DR Pfam: PF01425; Amidase; 1.
FT Hypochemical protein.
KW NON_TER
SQ SEQUENCE 151 AA; 16894 MW; D16F7383BADAA4B4D CRC64;

Query Match

Best Local Similarity 0.8%; Score 8; DB 4; Length 151;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 713 PLLLRL 720

Db 7 PLLLRL 14

RESULT 13

09F9V0 PRELIMINARY; PRT; 154 AA.
AC 09F9V0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Paal.
GN PAAL.
OS Azococcus evansii.
OC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group;
OC Azococcus.
OX NCBI_TaxID=59406;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KB740;
RA MEDLINE-20096669; PubMed-10629172;
RA El-Said Mohamed M.;
RT "Biochemical and molecular characterization of phenylacetate-coenzyme A ligase, an enzyme catalyzing the first step in aerobic metabolism of phenylacetic acid in Azococcus evansii.";
RL J. Bacteriol. 182:286-294(2000).
DR EMBL: AF176259; AAC28667.1;
DR InterPro: IPR003736; DUF157.
DR Pfam: PF02584; DUF157; 1.
DR TIGRfams: TIGR00369; unchar_dom_1; 1.
SQ SEQUENCE 154 AA; 16550 MW; 030AC598B5A441C3 CRC64;

Query Match

Best Local Similarity 0.8%; Score 8; DB 2; Length 154;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 681 FITLADT 688

Db 69 FITLADT 76

RESULT 14

09XTY5 PRELIMINARY; PRT; 157 AA.
AC 09XTY5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE T03D8.2 protein.
GN T03D8.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Mortimore B.J.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE-99069613; PubMed-9851916;
RA none;
RT "genome sequence of the nematode C. elegans: A platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z92838; CAB07406.1;
DR InterPro: IPR000230; Ribosomal_S12.
DR Pfam: PF00164; Ribosomal_S12; 1.
DR PRINTS: PR01034; Ribosomal_S12.
DR PRODOM: PD000576; Ribosomal_S12; 1.
DR TIGRfams: TIGR00961; TpsL_bact; 1.
SQ SEQUENCE 157 AA; 17370 MW; 6C6A837282EB26C CRC64;

Query Match

Best Local Similarity 0.8%; Score 8; DB 5; Length 157;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 501 TLFTHQSS 508

Db 27 TLFTHQSS 34

RESULT 15

09VRE4 PRELIMINARY; PRT; 207 AA.
AC 09VRE4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG1739 protein.
GN CG1739.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eurygaster; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahiye C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Nobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003569; AAF50857.1; -
 DR FlyBase: FBgn0031152; CG1739.
 SO SEQUENCE 207 AA; 23166 MW; 2AC6AC1BB60EA197 CRC64;

Query Match 0.8%; Score 8; DB 5; Length 207;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 LQEEEPPLA 434
 |||||||||
 DB 145 LQEEEPPLA 152

Search completed: May 7, 2003, 19:45:51
 Job time : 55 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 13, 2003, 21:52:01 ; Search time 525 Seconds

(without alignments)
4413.917 Million cell updates/sec

Title: US-09-929-769-7

Sequence: 1 MHLLVHAMVITLLCPRA.....MDPSAQISEALRIILHEAVM 1029

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV-xlh
-O/cgml_1/USPTO.spool/US09929769/runat_07052003_161923_8024/app_query.fasta.1.1223
-DB-nucleicseq_101002 -OPMT-fastlap -SUFFIX-p2n.rng -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIG=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT-pco -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR-US09929769.ecgn.1.1.34.4 @runat_07052003.161923.8024 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGOUT -DEV.TIMEOUT=120
-MAIN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5241	100.0	3437	21 AAC58109	Human PRO3434 nucl
2	5241	100.0	3437	21 AAA51266	Human DNA encoding
3	5241	100.0	3437	24 ABK11743	DNA encoding tumou
4	4211	80.3	2824	24 ABR35138	Human cDNA encodin
5	3229	61.6	2275	21 AAC76704	Human ORFX ORF2259
6	2055	39.2	1559	21 AAZ98062	Human secreted pro
7	2055	39.2	1559	22 AAAD11675	Human secreted pro
8	2055	39.2	1559	24 ABR69771	Human secreted pro
9	924	17.6	639	20 ABR86653	EST clone C2496.
10	891	17.0	526	22 ABA08362	Human secreted pro
11	863	16.5	6207	23 ABLO5745	Drosophila melanog
12	863	16.5	8355	23 ABLO5744	Drosophila melanog
13	592	11.3	484	22 AAL01467	Human reproductive
14	228	4.4	4351	23 ABLO7340	Drosophila melanog
15	193	3.7	119	21 AAC11139	Human secreted pro
16	163.5	3.1	14800	24 ABLE6291	Lung cancer relate
17	157.5	3.0	7419	23 AAS51427	Pseudomonas aerugi
18	157	3.0	33529	23 AAS17367	DNA sequence of S.
19	156.5	3.0	49377	19 AAV05287	The soraphen biosy
20	153.5	2.9	14835	24 AAS94858	Human DNA sequence
21	152.5	2.9	4080	19 AAV34368	Human protein tyro
22	152.5	2.9	4080	20 AA211221	Human protein tyro
23	152.5	2.9	4080	24 ABLE4464	Stomach cancer rel
24	151.5	2.9	3787	22 ABA01228	Putative TBP inter
25	151.5	2.9	6045	22 AAH77994	Nucleotide sequenc
26	151.5	2.9	7093	24 AAD38851	Human kinase (PKIN
27	151	2.9	3250	23 AAS84219	DNA encoding novel
28	150.5	2.9	28958	18 AAR89956	Sorangium cellulos
29	150.5	2.9	28958	17 AAR75299	DNA sequence of So
30	150	2.9	28598	17 AAT06769	Sorangium cellulos
31	148	2.8	15355	22 AAB57405	Human skeletal mus
32	148	2.8	15359	22 ABA08724	Human ryanodine re
33	148	2.8	15359	22 AAS69156	DNA encoding novel
34	148	2.8	15464	22 AA011415	Ryanodine receptor
35	147	2.8	4031	20 AAV69717	Tumour rejection a
36	147	2.8	4121	24 ABR83720	Human cDNA differe
37	147	2.8	4121	24 ABLE2861	Breast cancer rela
38	147	2.8	4121	24 ABLE3082	DNA encoding novel
39	147	2.8	5526	23 AAS76845	Human polynucleoti
40	147	2.8	5778	22 AAK53086	Human cDNA encodin
41	147	2.8	5804	22 AAK52102	Human ORFX ORF94
42	147	2.8	11736	22 AAS22652	Human chromosome 1
43	146.5	2.8	4224	21 AAC74939	Tumour rejection a
44	145.5	2.8	4827	22 AAH22143	
45	145	2.8	4225	20 AAV69720	

ALIGNMENTS

RESULT 1

AAC58109 standard; cDNA: 3437 BP.

25-JAN-2001 (first entry)

Human PRO3434 nucleotide sequence SEQ ID NO:22.

Human: tumour; diagnosis; neoplastic disease; proliferation; cancer;

Identification; tumorigenesis; anticancer; detection; ss.

Homo sapiens.

WO200053750-A1.

14-SEP-2000.

PCR/US99/28551

XX 02-DEC-1999; 99WO-US28551.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 01-SEP-1999; 99WO-US20111.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28634.
XX
PA (GETH) GENENTECH INC.
PI Botstein D, Goddard A, Gurney AL, Roy MA, Watanabe CK, Wood WT;
XX WPI; 2000-594320/56.
DR P-PSDB; AAB24027.
XX
PT Antibodies specific for PRO polypeptides, used to diagnose and inhibit
PT the growth of tumors in mammals, and to identify inhibitors of PRO
PT polypeptide activity or expression -
PS
PS Claim 50; Fig 13; 226pp; English.

CC The present invention describes an antibody that binds to a human
CC protein (I) selected from: PRO381; PRO1269; PRO1410; PRO1755; PRO1780;
CC PRO434; PRO1927; PRO3567; PRO1293; PRO1303; PRO344; PRO454;
CC PRO4397; PRO4407; PRO1555; PRO2038; and PRO2262. (I) has
CC anticancer activity and can be used to diagnose tumors in mammals, by
CC detecting complex formation when the antibody (I) is contacted with test
CC cells. Increased expression of genes encoding (I) can also be detected
CC to diagnose tumors. Agents which inhibit the activity of (I),
CC especially the antibodies, or an antisense oligonucleotide which
CC hybridizes to genes encoding (I), can be used to inhibit tumor growth,
CC preferably by inducing cell death. Methods from the present invention
CC can be used to identify compounds which inhibit the biological activity
CC of (I). AAC58019 to AAC58102 represent PCR primers and hybridisation
CC probes used in examples from the present invention for human PRO
CC sequences. AAC58103 to AAC58122 and AAB24021 to AAB24040 represent human
CC PRO polynucleotide and protein sequences given in the exemplification of
CC the present invention.
XX
XX Sequence 3437 BP; 558 A; 1186 C; 1098 G; 595 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	3437
Score:	5241.00	Matches:	1029
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-929-769-7 (1-1029) x AAC58109 (1-3437)

QY 1 MethHisIleLeuValAlaHisAlaMetValIleLeuLeuThreLeuGlyProProGAla 20
DB 46 ATGCACATCTCTGCTGCTCCATGCATGATGATCTGCTGAGCCGCCCTCCAGCC 105
QY 21 AspAspSerGluPheGlnAlaLeuLeuAspIleTyrPheProGluGluysProLeuPro 40
DB 106 GACGACAGGAGATTCCAGCGCTGACATCTGTTCCGAGAGAGAACCCAGCCGCC 165
QY 41 ThrAlaPheLeuValAspThrSerGluGluAlaLeuLeuLeuProAspThrLeuLeu 60
DB 166 ACCGCTCTCTGCTGACACATCGAGAGAGCCCTGCTCTCTGCTGAGAGCTG 225
QY 61 ArgMetIleArgSerGluValLeuAlaArgLeuValAspAlaLeuGlnAspLeuGluPro 80
DB 226 CGCATATCGCTTCTGAGGCTCCGCCCTGCTGAGAGCCGCCCTGAGAGCTGAGCC 285
QY 81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyTyrProValSerSerMetSerLeu 100
DB 286 CAGCAGCTCTCTGCTGCTGAGAGCTGCTGAGATCCGCCCTGAGAGAGCAAACTC 345
QY 101 LeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGluGlnAsnIleMet 120

DB 346 CTCACATCTCTGAGCAGCAGGAGCCAGACCCCACTCTGAGAGCAACATCATG 405
QY 121 AspIleAsnTyrMetAlaHisLeuValGluValGlnHisGluArgIleAspGlyGly 140
DB 406 GACAAAGATTACATGAGCCACCTGTGTGAGATCTACATAGAGCGCCGCTCCGAGCC 465
QY 141 GlnThrPheHisSerLeuLeuThrAlaSerLeuProProArgArgAspSerThrGlnAla 160
DB 466 CAGACTTTCACCTCTTGGCTGACAGCCTCCCTGCGCCCGCGAGACAGCAGAGGCA 525
QY 161 ProLysProLysSerSerProGluGlnProIleGlyGlnGlyArgIleArgValGlyThr 180
DB 526 CCCAAACCAAGAGCAGCCAGCCAGCCATAGCCAGAGCCGAGTTCGGGGGAGACC 585
QY 181 GlnLeuArgValLeuGlnIleProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro 200
DB 586 CAGCTCCGGGTGCTGGCCCTGAGAGCAGCTGGGATGTCTCCAGATTTTCCCG 645
QY 201 LeuSerProAspProArgTrpGlnSerSerSerProArgProValAlaLeuAlaLeuGln 220
DB 646 CTCAGCCCGGAGCCCTGCGTGGAGAGCTCCAGTCCCGCCGCTGAGCCCTCCGAG 705
QY 221 GlnAlaLeuGlyGlnGlnLeuAlaArgValValGlnGlySerProGluValProGlyTyr 240
DB 706 CAGCCCTGAGCAGAGAGCTGAGCCCTGCTCCAGAGGAGCCAGAGCTCCGAGCCATC 765
QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyGlyAlaLeu 260
DB 766 ACGTGTGCTGTCTGAGAGCCCTCCGACCTCTCTAGCTCCCAACAGCGGTGCTG 825
QY 261 ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgIleLeuCysGln 280
DB 826 GTGATGTTCATGACGACCTGAGCCACTTCTGCGCTGCGCTGAGCCAGCTCTCCAG 885
QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuysValLeuLeu 300
DB 886 TACACAGGCTGTGTGACAGAGACCGGCTTCTCTGCTCTTCTGAGAGTGTCTCTG 945
QY 301 GlnMetLeuGlnIlePheLeuAspSerProGlyValGluGlyIleProLeuArgAlaGlnLeu 320
DB 946 CAGATGCTGAGTGGCTGAGCAGCCCTGAGGTGAGAGCGGCGCCCTGCGGAGCAGCTC 1005
QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyLeu 340
DB 1006 AGGATGCTTGCAGCCAGCCAGGCTCAGCCGCGAGGCTCAGTATGTGCGAGGGGCTC 1065
QY 341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGlnValValSerSerThrVal 360
DB 1066 CTGCGCTGAGCAGGCGCTGAGCTTCTGCTGAGGAGCTGAGGTGCTCAGCTCAGCCGTC 1125
QY 361 ArgAlaValIleAlaThrLeuArgSerGlyGlnGlnCysSerValGluProAspLeuIle 380
DB 1126 CGTGCCGTATCCCAACCTGAGAGTGTGGGAGACAGTGCAGTGCAGCGAGCTGATC 1185
QY 381 SerIleValLeuGlnGlnIleGlnValArgSerProHisLeuGlnGlnLeuLeuThr 400
DB 1186 AGCAAAAGTCTCCAGGCGCTGATGAGTATGCCCTCCAGTGAAGAGAGCTGTGACT 1245
QY 401 AlaPhePheSerAlaThrAlaAspAlaAspLeuProPheProAlaCysIleProValAla 420
DB 1246 GCATTTCTCTCTCCAGCTGAGGATGCTCTCCCTTTCCAGACCTGTAAAGCCGTTGTC 1305
QY 421 ValIleSerSerLeuLeuLeuGlnGlnGlnGlnIleProLeuAlaGlyIleysProGlyAla 440
DB 1306 GTGCTGAGCTCCCTGCTGCTGAGAGAGAGACCCCTGCTGAGGAGGAGCGGTGAG 1365
QY 441 AspGlyIleSerLeuGlnAlaValArgLeuGlyProSerSerSerIleLeuLeuValAspTrp 460
DB 1366 GACGTGTGAGCCTGAGAGCGGTGAGGCGGCTGCTGAGAGCCCTCTAGTGAAGCTG 1425
QY 461 LeuGlnMetLeuAspProGluValValSerSerCysProAspLeuGlnLeuArgLeuLeu 480

PR 12-JAN-1999; 9905-0115558.
 PR 12-JAN-1999; 9905-0115565.
 PR 12-JAN-1999; 9905-0115733.
 PR 09-FEB-1999; 9905-0119341.
 PR 10-FEB-1999; 9905-0119341.
 PR 12-FEB-1999; 9905-0119337.
 PR 02-JUN-1999; 9905-0119965.
 PR 02-JUN-1999; 9905-0122522.
 XX (GETH) GENENTECH INC.
 PA
 XX Boistein D, Desnoyers L, Ferrara N, Fong S, Gao W, Goddard A,
 PI Currey AL, Pan J, Roy MA, Stewart TA, Tumas D, Watanabe CK;
 PI Wood WL;
 XX WPI; 2000-431586/37.
 DR P-PSDB; AAY96736.
 XX
 PT Isolated nucleic acid molecule encodes a PRO polypeptide which is a
 PT transmembrane polypeptide
 XX
 PS Claim 12; Fig 15; 154pp; English.
 CC This cDNA encodes PRO3434, a novel secreted protein. The invention
 CC concerns novel secreted and transmembrane proteins, designated PRO
 CC polypeptides. The cDNA and gene sequences are useful in the recombinant
 CC production of PRO polypeptides, as a hybridization probe to screen
 CC libraries to isolate cDNAs with sequence identity to PRO polypeptides or
 CC to map the gene encoding the PRO polypeptides and analyzing genetic
 CC disorders. The cDNA/gene can also be used to produce transgenic animals
 CC useful for the development and screening of therapeutically useful
 CC reagents. They can also be used in gene therapy, e.g. to replace a
 CC defective gene.
 XX
 SQ Sequence 3437 BP; 558 A; 1186 C; 1098 G; 595 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3437
 Score: 5241.00 Matches: 1029
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-929-769-7 (1-1029) x AAA51266 (1-3437)

QY 141 GlnThrPheHisSerLeuLeuThraLeuProProArgArgSerThrGluAla 160
 DB 466 GACATTTTCCACTCTTGTCTCACAGCCCTCCCTGCGCCCGGAGACAGCAGAGCA 525
 QY 161 ProGlyProLysSerSerProGluGlnProIleGlyGlnGlyArgIleArgValGlyThr 180
 DB 526 CCCAAACCAAGAGCAGCCAGCAGCCATAGAGCCAGGCGGAGGCTTGGGGGGAGCC 585
 QY 181 GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro 200
 DB 586 CAGCTCCGGGCTCTGGGCGCCCTGAGGACGACGACGTCGTCGATGCTCCAGATTTTCCGG 645
 QY 201 LeuSerProAspProArgTrpIleSerSerSerProArgProValAlaLeuAlaLeuGln 220
 DB 646 CTCAGCCCGGACCTCTGGTGGCAGAGCTCCAGTCCCGCCCGCTGCGCCCTCCGACG 705
 QY 221 GlnAlaLeuGlyGlnGluLeuAlaArgValValGlnGlySerProGluValProGlyIle 240
 DB 706 CAGGCGCTGGCCAGAGAGCTGGCCCGCTGTCAGGGGAGCCCGCAGGTCCGGGAGATC 765
 QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyAlaLeu 260
 DB 766 AGGTCGTCGTCTCTGACGGCCCTCGCCACCTGCTCAGCTCCCGACAGCGGCTGGCTTG 825
 QY 261 ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGln 280
 DB 826 GTGATGTCATGACACCGTGGCCACTTCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
 QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeuLeu 300
 DB 886 TACACGCGCTGTCTGTCACAGGACCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 945
 QY 301 GlnMetLeuGlnTrpLeuAspSerProGlyValGlnGlyProLeuArgAlaGlnLeu 320
 DB 946 CAGATGCTCAGAGCTGTGACAGCCCTGCTGAGGGGCGGCGCCCTCGGCGACAGCTTC 1005
 QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeu 340
 DB 1006 AGGATCTTGGCAGCCAGGCTCTGACGGGCGGAGCTGAGTATGTGGAGGGGGCTTC 1065
 QY 341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGlnValValSerSerThrVal 360
 DB 1066 CTGCGCTGGCCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
 QY 361 ArgAlaValIleAlaThrLeuArgSerGlyGlnCysSerValGlnProAspLeuIle 380
 DB 1126 CGTCCGTCATGCGCCACCTGAGGTCGTGGGAGCACTGACGCTGAGCGCGCTGATC 1185
 QY 381 SerIleValLeuGlnGlyLeuIleGluValArgSerProHisLeuGlnLeuLeuThr 400
 DB 1186 AGCAAAAGTCTCCAGGGGCTGATCGAGGTGAGGTCCCGCCACCTGAGGAGCTGCTGACT 1245
 QY 401 AlaPheSerAlaThrAlaAspAlaAspProPheProAlaCysLysProValVal 420
 DB 1246 GCATTTTCTTGCCACTCGGATGCTGCTCCCGCTTTCAGAGCTGTAAAGCCGCTTGG 1305
 QY 421 ValValSerSerLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 440
 DB 1306 GTGTGAGCTCTCTGCTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1365
 QY 441 AspGlyGlySerLeuGlnValAlaArgLeuGlyProSerSerGlyLeuLeuValAspTrp 460
 DB 1366 GAGGTGGAGGCTGGAGGCGCTGCGGCTGGGGGCTCTGCAAGGCTCTAGTGAATG 1425
 QY 461 LeuGlnMetLeuAspProGluValValSerSerCysProAspLeuGlnLeuArgLeuLeu 480
 DB 1426 CTGGAAATGCTGGAGCCGAGGTGTCAGACAGTGGCCGAGCTGAGCTGAGCTGCTG 1485
 QY 481 PheSerArgArgGlyGlyGlyGlyGlnAlaGlnValProSerPheArgProGlyLeuLeu 500
 DB 1486 TTCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1545
 QY 501 ThrLeuPheThrHisIleSerIleSerTrpProThrIleHisGlnCysIleArgValLeuLeu 520

Db 1546 ACCCTCTTCAGCATGCTCCAGCTGGCCACATGACCCAGCTGATCCGACTCTGCTG 1605
 Qy 521 GlysSerArgGluGlnArgPheAspProSerLaserLeuAspPheLeuTrpAlaGlu 540
 Db 1606 GGCAGAGCCGGGAGACAGAGGTTCAGCCCTCTGCTCTGAGACTCTCTCTGGGCGTGC 1665
 Qy 541 IleHisValProArgIleTrpGlnGlyArgAspGlnArgThrProGlnLysArgArgGlu 560
 Db 1666 ATCATGATCTCTCGCATCTGCGACAGGGGGGAGCACCGCACCCGACAGACCGCGGAG 1725
 Qy 561 GluLeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGluLeuIleLeuAla 580
 Db 1726 GAGCTGTGCTGGCGGCTCCAGGGCCCGAGCTCATCCCTGCTGAGACTGATCTGTGACC 1785
 Qy 581 GluAlaGluThrArgSerGlnAspGlyAspThrAlaIleCysSerLeuIleGlnAlaIleArg 600
 Db 1786 GAGCGGAGAGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1845
 Qy 601 LeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValIleHisGlu 620
 Db 1846 CTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1905
 Qy 621 HisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgArgCysArgAsp 640
 Db 1906 CACCTGTACAGCTCATACACAGCTGGGAGACAGCTGTGGAGAGGCTGCGCCGAGAC 1965
 Qy 641 LeuLeuLeuGlnLeuLeuTrpGlnArgProGluLeuArgValProValProGluValLeu 660
 Db 1966 CTCTCTCTGAGCTCTACCTACACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2025
 Qy 661 LeuHisSerGlnGlyAlaAlaSerSerSerValCysLysLeuAspGlyLeuIleHisArg 680
 Db 2026 CTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2085
 Qy 681 PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGluAsnArgGlyAlaAsp 700
 Db 2086 TTCATTCACGCTCTTGGGAGACACAGGAGCTCCGCGGCTGGAGAACCGGAGGCGGAT 2145
 Qy 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuLeuArgHisLeu 720
 Db 2146 GCCAGCATGGCTCTCCGAGACTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2205
 Qy 721 ProMetIleAlaAlaLeuLeuHisGlyArgThrHisLeuAsnProGlnGluPheArgGln 740
 Db 2206 CCATGATCGGGGCGCTCTCCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2265
 Qy 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGluLeuLeuLeuProHis 760
 Db 2266 CAGAACACCTGAGCTCTCTCCAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 2325
 Qy 761 ValPheArgSerGlnHisGlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeu 780
 Db 2326 GTGTTCGCGAGCGAGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2385
 Qy 781 LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln 800
 Db 2386 CTGCTGAATTAACAGAGAGTCTCTCCGCGCATGTGGCTGCTCATCAACAAGTTTGTGCAG 2445
 Qy 801 PheIleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHis 820
 Db 2446 TTCATTCATTAAGTATACCTACCAATGAGCCACAGCACTCTCTCTCCAGAGAGCAG 2505
 Qy 821 AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu 840
 Db 2506 GCCGACCCGCTCCAGAGCTGCTCTTGCACACAGTGAAGCTGGAGAGCTGAATTCCTCTC 2565
 Qy 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgThrAspArgGlyLeuAspGluGlu 860
 Db 2566 CTGGCAGAGGCTCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2625
 Qy 861 GlyGluGluLeuSerSerArgAlaGlySerLeuProLeuValSerValSerLeuPheThrPro 880

Db 2626 GCGGAGAGGAGAGCTCAGCCGCGCTCTTCCCTGCTGAGGCTCTCCCTTTACACCCCT 2685
 Qy 881 LeuThrAlaAlaGluMetAlaProTyrMetLysArgLeuSerArgGlyGlnThrValGlu 900
 Db 2686 CTGACCGCGCGCGAGATGGCCCCCTTACATGAAAGGCTTTCCCGGAGCCAAAGCGGTGAG 2745
 Qy 901 AspLeuLeuGluValLeuSerAspIleLeuSerAspGlnMetSerArgArgArgProGluIleLeu 920
 Db 2746 GATCTGTGAGAGGTCTGAGTGACATPAGACAGATAGTCCCGGAGAGACCGAGATCTCTG 2805
 Qy 921 SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGluGluCysArgAsn 940
 Db 2806 AGCTTCTTCTGACCACTGACGCGGCTGATGAGACTGCGCGAGAGAGTGTGGCGCAAC 2865
 Qy 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaIleAlaPhe 960
 Db 2866 CTGCGCTTACAGCTCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2925
 Qy 961 LeuProThrPheMetTyrCysLeuGlySerGlnAspPheGluValGlnThrAlaLeu 980
 Db 2926 CTGCGCAGGTTTCACTTACTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2985
 Qy 981 ArgAsnLeuProGluTrpAlaLeuLeuLeuGluGlnGlnHisAlaAlaValLeuLeuHisArg 1000
 Db 2986 CGGAACCTGCTGAGTACGCTCTCTGTCGCAAGAGACGCGGCTGTGCTCTCCACCGG 3045
 Qy 1001 AlaPheLeuValGlyMetTyrGlyGlnMetAspProSerAlaGlnIleSerLysAlaLeu 1020
 Db 3046 GCCTTCTGTGGTGGCATGACGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3105
 Qy 1021 ArgIleLeuHisMetGluAlaValMet 1029
 Db 3106 AGGATCTCGCATATGAGAGCGCTGATG 3132

RESULT 3
 ID ABK11743 standard; cDNA: 3437 BP.
 AC ABK11743;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE DNA encoding tumour associated antigenic target polypeptide (TAT) 156.
 KW Tumour associated antigenic target polypeptide; TAT; cancer;
 KW breast cancer; colorectal cancer; lung cancer; ovarian cancer;
 KW central nervous system cancer; liver cancer; bladder cancer;
 KW pancreatic cancer; cervical cancer; melanoma; leukemia; TAT156;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH key Location/Qualifiers
 FT sig_peptide 86..93
 FT CDS 86..3135
 FT /tag= a
 FT /tag= b
 FT /product= "TAT156"
 FT /note= "Tumour associated antigenic target
 mat_peptide 94..3132
 FT /tag= c
 FT /label= Mature_TAT156
 XX
 XX W0200216581-72. PCT/us01/25464 (5/14/01)
 XX 28-FEB-2002.
 XX 14-AUG-2001; 2001WO-US25464.
 XX 24-AUG-2000; 2000WO-US23328.
 XX 28-FEB-2001; 2001WO-US06520.
 PR 22-JUN-2001; 2001US-0888257.

Db 1546 ACCCTCTTACAGCATCATCCAGCTGGCCACACATGCATCCAGTCCGTCTG 1605
 QY 521 GLYSSERARGGLUGLARGPHEASPROSERALASERLEUASPHELEUTRPLACY 540
 Db 1606 GGCAGACCCGGGAAACAGAGGTTCACCCCTCTGCTCTGAGACTTCTGGGGCTGC 1665
 QY 541 ILEHISVALPROARGILETRPGLNGLYARGASPGINACTFTHPROGLNLSARGTRGL 560
 Db 1666 ATTCATGTTCTCGCATCTGGCAGGGGGGGGAGCCAGCCACCCCGGAGAGGGGGAG 1725
 QY 561 GLULEUVALLEUARGVALGLNGLYPROGLULEUILESERLEUVALGLULEULEUALA 580
 Db 1726 GAGCTGGTGTCTGGCTCCAGGGCCCGGAGCTCATCAGCTGGTGGAGCTGATCTCTGCC 1785
 QY 581 GLULIAGLUTRARGSERGLASPGILYSPHRAALACYSSELEULEGILNGLARG 600
 Db 1786 GAGGGGGAGAGCGGAGCGGAGGAGCCAGCCGCTCAGGCTCATCCAGGCCCGG 1845
 QY 601 LEUPROLEULEULEUSERCYSYSGLYASPAJUSERVALARGLYVALTRGLU 620
 Db 1846 CTGCCCTCGTCTAGCTGCTGCTGGGAGCATGAGAGTCTCAGAGAGGTGACGGAG 1905
 QY 621 HISLEUSERGLYCYSLIENGLINTRPGLYASPSERVALLEUGLYARGARGCYSARGSP 640
 Db 1906 CACCTGTGAGCTGCATCCAGAGTGGGGAGACACGCTGTGGGAAGCGCTGCCGAGAC 1965
 QY 641 LEULEULEUGLINTLEUTYLEUGLINTPRGILULEARGVALPROVALPROGLUVALLEU 660
 Db 1966 CTTCCTCGCAGCTCTACTCTACGAGCGGAGCTGGGGGCTGCCGTGCTGAGGTCTCA 2025
 QY 661 LEUHSISERGLUGLYALALASERSESERVALCYSLYLEUASPGIYLEUILEHSARG 680
 Db 2026 CTGCACACGGAAGGGGCTGCCAGCAGACGCTGCAAGCTGGAGGACTCATCCAGCGC 2085
 QY 681 PHEILETRHLEULEULEALASPTHISERASPSERARGALALEUGLUNASARGLYALASP 700
 Db 2086 TTCAATCAGCTCTCTTCCGACACACGACATCCCGGGCTGGAGAACCGAGGGCGGAT 2145
 QY 701 ALASERMETALACYARGLYSLEUALAVALALAHISPROLEULEULEUARRGHISLEU 720
 Db 2146 GCGACATAGGCTGCGGAGAGCTGGGCTGGCGCACCCGCTCTCTGCTCAGGCACTCG 2205
 QY 721 PROMETILEALALALEULEUHSISGLYARGTRHISLEUASNPHENGLINUPHEARGLIN 740
 Db 2206 CCCAATGATCGGGGCTCTCTGACAGCGGACCCACCTCACTTCCAGAGATTCCGGCAG 2265
 QY 741 GLINASNHSLEUSERCYPHELEUHSISVALLEUGLYLEULEUGLULEULEUNPROHS 760
 Db 2266 CAGAACCACTGAGCTGTCTCTGACAGCTGGGGCTGTGGAGCTGCTGCAGCCGAC 2325
 QY 761 VALPHEATGSEGLUHSISGLNGLYALALEUTRPAJPCYSLLEULEUSERPHEILEARGLEU 780
 Db 2326 GTTGTCCGACGAGCACAGGGGGGCTGTGGAGCTGCTCTCTCTCTCTCTCTCTCTG 2385
 QY 781 LEULEUASNTRYARGLYSSESERARGHISLEUALAALAPHEILEASNLSPHEVALGLIN 800
 Db 2386 CTGCTGAATACAGGAATCTCTCCGCCATCTGCTGCTTATCAACAAGTTTGACAG 2445
 QY 801 PHEILEHSISLYSTRILETHRTYRASNALAPROALALALILESERPHELEUCLINYSHS 820
 Db 2446 TTCAATCCATTAAGTACATTACATCAATGCCCGAGCAGCATCTCTCTCTGCAAGAGC 2505
 QY 821 ALAASPPROLEUHSISASPLEUSERPHEASPSANSERASPLEUVALMETLEUYSSELEU 840
 Db 2506 GCGGACCGGCTCCAGACTGTCTTCGACAAACATGACTGTGTAATGTAATCCCTC 2565
 QY 841 LEUALAGLYLEUSERLEUPROSERARGASPARGTHTASPARGLYLEUASPGIUGLU 860
 Db 2566 CTTCGAGAGGCTCAGCCTCCAGCAGGAGCAGACGAGCAGGAGCCTGAGCAAGAG 2625
 QY 861 GLYGLUGLUGUSERSERALAGLYSERLEUPROLEUVALSERVALSERLEUPHERHPRO 880
 Db 2626 GCGGAGGAGAGGCTCAGCGGGCTCTTGGCCCTGTGCTGCTCTCTCTCTCTCTCTCT 2685

QY 881 LEUTHRALALAGLUMETALAPROTYRMETLYSARGLEUSERARGLYGLINTHRVALGLU 900
 Db 2686 CTGACCGGCGGCGAGATGGCCCCCTTACATGAACGGCTTTCCCGGGCCAAACGGTGGAG 2745
 QY 901 ASPLEULEUGLUALLEUSERASPILEASPGIUMETSERARGARGARPROGLULEULEU 920
 Db 2746 GATCTGTGGAGGTTCTGAGTACATPAGACGATGTCCCGGGAGACCCAGATCTCG 2805
 QY 921 SERPHEPSESERTHASNLEUGLINARGLEUMETSERSERALAGLUGLYCYSARGASN 940
 Db 2806 AGCTTCTTCTCGACCAACCTGCGAGCGGTGATGAGACTCGGGCCAGAGAGTGTGGCAAC 2865
 QY 941 LEUALAPHESEERLEUALLEUARGSERMETGLINASNPROSERILEALALALAPHE 960
 Db 2866 CTGCGCTTACAGCTGGCCCTCGGCTCTCATGCAAAACGCCCCAGCATTCACGCCCTTTC 2925
 QY 961 LEUPROTHRPHMETLYTCYSLLEUGLYSERGLINASPHEGLUVALGLINTHRALALEU 980
 Db 2926 CTGCCACGTTGATGACTGCTGGGAGCCAGGACTTTGAGGTGTGGCAGACGGCCCTC 2985
 QY 981 ARGASNLEUPROGLUTYRALALEULEUCYSGINGLUNHISALALAVALLEULEUHSARG 1000
 Db 2986 CGGAACCTGCGAGTACGCTCTCTGTGCCAAAGACACGCGCTGTCTGCCACCGG 3045
 QY 1001 ALAPHELEUVALGLYMETYRGLYGLINMETASPPROSERALAGLILESERGLUALLEU 1020
 Db 3046 GCTTCTCTGTGGGTGATGATGAGGCGGACGATGAGCCCGAGCGCCAGATCTCCAGGCCCTG 3105
 QY 1021 ARGILELEUHSMETGLUALAVALEMET 1029
 Db 3106 AGCATCTGATATGAGAGCGCTGATG 3132
 RESULT 4
 ID ABK35138 standard; cDNA; 2824 BP.
 AC ABK35138;
 XX
 DT 08-MAY-2002 (first entry)
 DE Human cDNA encoding secreted protein #276.
 KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
 KW viral infection; bacterial infection; fungal infection; diabetes; asthma;
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
 KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
 KW tissue regeneration; wound healing; burn; haematopoiesis;
 KW myeloid cell deficiency; lymphoid cell deficiency.
 OS Homo sapiens.
 PN WO200177288-A2.
 PD 18-OCT-2001.
 PF 29-MAR-2001; 2001WO-US10224.
 PR 06-APR-2000; 2000US-195582P.
 PA (GENY) GENETICS INST INC.
 PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 DR Gulukota K, Graham JR;
 DR WPI: 2002-179321/23.
 PT Five hundred and ninety two polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for
 PT treating immune deficiencies and disorders such as autoimmune disorders

XX Claim 1, Page 210, 372pp: English.

CC The invention relates to 592 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins. The polynucleotides can be used as probes for the
 CC identification and isolation of full length cDNA and genomic DNA. The
 CC polynucleotides and proteins can also be used as nutritional supplements.
 CC The proteins are useful in the treatment of various immune deficiencies
 CC and disorders such as viral infections, bacterial infections, fungal
 CC infections, autoimmune thyroiditis (e.g. rheumatoid arthritis, multiple
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
 CC and conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment of
 CC burns, incisions and ulcers. The proteins are also useful for regulating
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
 CC Sequences ABK34853-ABK35454 represent polynucleotides of the invention.

XX Sequence 2824 BP; 443 A; 982 C; 910 G; 489 T; 0 other;

Alignment Scores:

Pred. No.:	3,3e-302	Length:	2824
Score:	4211.00	Matches:	829
Percent similarity:	98.81%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	0
Query Match:	80.35%	Indels:	10
DB:	24	Gaps:	1

US-09-929-769-7 (1-1029) x ABK35138 (1-2824)

QY 200 ProLeuSerProAspProArgTrpGlnSerSerProArgProValAlaLeuAlaLeu 219
 DB 2 CGGCTCAGCCCGGAGCCCTGGTGGAGAGCTCCAGTCCCGCCCGCGGCTTCGCCCTG 61
 QY 220 GlnGlnAlaLeuGlnGlnGlnLeuAlaArgValAlaGlnGlnSerProGlnValProGln 239
 DB 62 CACGAGGCGCTGGGCGAGGAGTGGCCCGCTGCTCCAGGAGCCCGCGAGTGGCGGC 121
 QY 240 IleThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlnValAla 259
 DB 122 ATACGGTGGGTGTCTCTGCGAGGCTGGCCACCTGCTGAGTCCGCCACAGGCGGTGGC 181
 QY 260 LeuValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuCys 279
 DB 182 CTGGTGTATGTCAMGACCGCTAGCCACTTCTGGCGTGGCGGCTGGCGGCGAGCTGTC 241
 QY 280 GlnTyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeu 299
 DB 242 CAGTACAGGCGTGTGTGGCCAGACACCGGCTTCTCTCCGCTCTCCGAGGCTCTC 301
 QY 300 LeuGlnMetLeuGlnTrpLeuAspSerProGlnValGlnGlnGlnProLeuArgAlaGln 319
 DB 302 CTGCAATGTGTGAGTGGCGGACAGCCCTGGGTGGAGGGCGGCGGCGGCGGCGAG 361
 QY 320 LeuArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGly 339
 DB 362 CTCAGATGCTTCCAGCCAGGCGCTCAGCGGCGGCGGCGGCTGAGTGTGCGAGGGGG 421
 QY 340 LeuLeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGlnValAlaSerSerThr 359
 DB 422 CTCTCTCGCTGGCGGAGGCGCTGGCTTCCGTCAGAGCTGGAGGAGGTCACCTCCAC 481
 QY 360 ValArgAlaValIleAlaThrLeuArgSerGlyGlnGlnCysSerValGlnProAspLeu 379
 DB 482 GTCCGTCGCTCATTCGCCACCTGAGGTCTGGGAGACAGTGCAGCTGGAGCGGAGCTG 541
 QY 380 IleSerLysValLeuGlnGlnGlnLeuIleGlnValArgSerProHisLeuGlnGlnLeu 399
 DB 542 ATCAGCAAGTCTCTCAGGCGGTGATGAGTGAAGTCCCGCCACCGAGGAGGAGCTCTG 601

QY 400 ThrAlaPheSerSerAlaThrAlaAspAlaSerProPheProAlaCysLysProVal 419
 DB 602 ACTGATTTCTTCTCTGCGCATGCGGATGCTGCTCCCGCTTCCAGGCTGTAGCCCGGT 661
 QY 420 ValValAlaSerSerLeuLeuGlnGlnGlnGlnProLeuAlaGlyLysProGln 439
 DB 662 GTGGTGGTGGCTCTCCGCTGCTGTCAGAGAGAGAGCCCTGGCTGGGGGAGAGCCGGT 721
 QY 440 AlaAspGlyLysSerLeuGlnAlaValAlaArgLeuGlyProSerSerGlyLeuLeuValAsp 459
 DB 722 CGGAGAGGTGGCAGCGCGGAGGCGCTGGCGGCGGCGGCGGCTGTCAGGCTCTAGTGAG 781
 QY 460 TrpLeuGlnMetLeuAspProGlnValAlaSerSerCysProAspLeuGlnLeuArgLeu 479
 DB 782 TGGCTGGAATGCTGGAGCCCGGAGGTGTGTCAGAGCTGCCCGCCGACGCTCAGGCTG 841
 QY 480 LeuPheSerArgTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 499
 DB 842 CTCTTCTCCGAG 901
 QY 500 LeuThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeu 519
 DB 902 CTGACCTCTTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961
 QY 520 LeuGlyLysSerArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 539
 DB 962 CTGGGCAAGAGCGCGGAG 1021
 QY 540 CysIleHisValProArgIleTrpGlnGlyArgAspGlnArgTrpProGlnLysArgArg 559
 DB 1022 TGCATTCATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1081
 QY 560 GlnGlnLeuValLeuArgValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 579
 DB 1082 GAGAGAGTGGTGTGGCGGAG 1141
 QY 580 AlaGlnAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAla 599
 DB 1142 GCCGAGCGGAG 1201
 QY 600 ArgLeuPheLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysValThr 619
 DB 1202 CGGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1261
 QY 620 GlnHisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlnArgArgCysArg 639
 DB 1262 GAGACCTGTGACGCTCATTCAGACAGAGTGGAGAGACAGCTGCTGGCGAGGCGCTGCGA 1321
 QY 640 AspLeuLeuLeuGlnLeuTrpLeuGlnArgProGlnLeuArgValProValProGlnVal 659
 DB 1322 GACCTTCTCTGAGCTCTACCTTACACAGCGGCGGAGCTGGGCTGGCTGAGAGTCT 1381
 QY 660 LeuLeuHisSerGlnGlnAlaAlaSerSerSerValCysLysLysLeuAspGlyLeuHis 679
 DB 1382 CTACTGACAG 1441
 QY 680 ArgPheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnAsnArgGlyAla 699
 DB 1442 CGCTTCATCAGCTCTCTGCGGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
 QY 700 AsnAlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuLeuArgHis 719
 DB 1502 GATGCGAGCATGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1561
 QY 720 LeuProMetIleAlaAlaLeuLeuHisGlyArgThrHisLeuAsnPheGlnIlePheArg 739
 DB 1562 CTGCCATGATGCGGCGGCTCTGCAAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
 QY 740 GlnGlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuGlnPro 759
 DB 1622 CAGCAGAACACCTGAGAGTCTCTGTCAGAGTGTGAGGCTGTGGAGGCTGTGAGAGCGG 1681
 QY 760 HisValPheArgSerGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 779

```

|||||
Db 1682 CACGTGTTCCGACGAGACACAGGGGGCGCTGTGGGACCTCTTCTCTCATCCG 1741
Qy 780 LeuleuLeuSntYrArglySerSerArgHisLeuAlaIleAsnLysPheVal 799
Db 1742 CTGCTGTGAATTCAGAGAAATCTCCCGCATCTGCTGCTTCATCAACAGTTGTG 1801
Qy 800 GluPheIleHisLysTyrIleThrTyrAsnAlaProAlaIleSerPheLeuGlnLys 819
Db 1802 CAGTTTCATCCATTAAGTACATTACTACATGCCAGAGGACCATCTCTTCTGACAGAG 1861
Qy 820 HisAlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSer 839
Db 1862 CACGCCGACCGCTCCACGACCTGCTTCCGACAAACAGTGCATGATGATGCTGAATCC 1921
Qy 840 LeuleuAlaGlyLeuSerLeuProSerArgAspAspArgThrAspArgGlyLeuAspGlu 859
Db 1922 CTCTCTGAGGGCTCAGGCTCTCCAGCAGGAGGAGACAGACCGAGGCTGAGCAGAA 1981
Qy 860 GluGlyGluGluGluSerSerArgLysSerLeuProLeuValSerValSerLeuPheThr 879
Db 1982 GAGGGCCAGGAGAGAGCTCAGCGGCTCTGCCCCCTGGTCAAGCTCTCCCTGTTCAAC 2041
Qy 880 ProLeuThrAlaAlaGluMetAlaProTyrMetLysArgLeuSerArgGlyGlnThrVal 899
Db 2042 CCTCTGACCGCGCGGAGATGGCCCTTACATGAAGCGCTTCCCGGGGCCAAACGGGTG 2101
Qy 900 Glu-----AspLeuLeuGluValLeuSerAspIle 909
Db 2102 GAGGGTAGTCAGGCGCTCTTACCCAGCCAGATCTGCTGGAGAGTTCTAGTGACATTA 2161
Qy 910 AspGluMetSerArgArgArgProGluIleLeuSerPhePheSerThrAsnLeuGlnArg 929
Db 2162 GAGCAGATGTCCTCCGGGAGAGACCGAGATCCTGAGCTTCTCTGACCAACCTGAGGG 2221
Qy 930 LeuMetSerSerAlaGluGluCysArgAsnLeuAlaPheSerLeuAlaLeuArgSer 949
Db 2222 CTGATGAGCTCGCGGAGAGAGTGTGCGCGCAACCGCTTCAGCGCTGCGGCTGCC 2281
Qy 950 MetGlnAsnSerProSerIleAlaAlaIlePheLeuProIlePheMetTyrCysLeuGly 969
Db 2282 ATGCAGAACACCCCGACATTCAGACCGCTTCTGCCCCAGCTGATATGACGCTGGGC 2341
Qy 970 SerGlnAspPheGluValAlaGlnThrAlaLeuArgAsnLeuProGluTyrAlaLeuLeu 989
Db 2342 AGCCAGAGACTTGAAGTGTGTCAGACGGCCCTCGGAAACGCTGAGATGAGCTTCTG 2401
Qy 990 CysGlnGluHisAlaAlaValLeuLeuHisArgAlaPheLeuValGlyMetTyrGlyGln 1009
Db 2402 TGCCAGAGACACGGGCTGTGCTGCTCCACCGGGCTTCCTGCTGGCATGACGGCCAG 2461
Qy 1010 MetAspProSerAlaGlnIleSerGlnAlaLeuArgIleLeuHisMetGlnAlaAlaMet 1029
Db 2462 ATGGACCCAGCGGCGAGATCTCCGAGGCCCTGAGATCTGCAATATGAGAGCCGTGATG 2521

RESULT 5
AAC76704
ID AAC76704 standard; cDNA; 2275 bp.
XX
AC AAC76704;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2259 polynucleotide sequence SEQ ID NO:4517.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnereary; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antiinflammatory;
KW hypotensive; dermatological; immunosuppressive; antineuritic; antithyroid;
KW antiviral; antibacterial; antifungal; antineuritic; antithyroid;
KW antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

```

```

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
EN W0200058473-A2.
XX
PD 05-OCT-2000.
XX
PE 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinketsu RA, Leach M;
XX
DR WPI; 2000-602362/57.
XX
P-PSDB; AAB42495.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 5; Page 3715-3716; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnereary;
CC antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antineuritic;
CC antithyroid; and antinaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 2275 bp; 393 A; 766 C; 719 G; 397 T; 0 other.

Alignment Scores:
Pred. No.: 1,43e-229 Length: 2275
Score: 3229.00 Matches: 633
Percent Similarity: 98.44% Conservative: 0
Best Local Similarity: 98.44% Mismatches: 0
Query Match: 61.61% Indels: 10
DB: 21 Gaps: 1

US-09-929-769-7 (1-1029) x AAC76704 (1-2275)
Qy 397 GluLeuLeuThrAlaPheSerSerArgHisLeuAlaIleAsnLysPheVal 416
Db 22 GAGCTGCGACGAGATCTCTCTGCTCCAGTGGAGTGTGCTCCCGCTTCCACCGTGT 81
Qy 417 LysProValValValValSerSerLeuLeuGlnGluGluGluProLeuAlaGlyGly 436

```

Db 82 AACCCCTGTGTGGTGGAGCTCCCTGCTGCTCAGAGAGAGAGACCCCTGGCTGGGGG 141
 Qy 437 LysProGlyAlaAspGlyGlySerLeuGluAlaValArgLeuGlyProSerSerGlyLeu 456
 Db 142 AACCCGGGTGGAGCGGTGGCAGCTGGAGCGCGGTGGGGCTGGCCCTGCTGAGGCTC 201
 Qy 457 LeuValAspTrpLeuGlnMetLeuAspProGluValAlaSerSerCysProAspLeuGln 476
 Db 202 CTGAGTGGAGCTGGCAAAAGCTGAGACCCCGAGAGGTGGTCAAGCAGCTCCCGCAGCTG 261
 Qy 477 LeuArgLeuLeuPheSerArgArgGlyGlyGluAlaGlnAlaGlnValProSerPheArg 496
 Db 262 CTGAGGCTGCTCTTCTCCGAGAGAGGCAAAAGTCAAGGCCAGAGTCCCTGGTCCGT 321
 Qy 497 ProTyrLeuLeuThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIle 516
 Db 322 CCTACTCTCTGACCTCTTCAAGCATCAGTCCAGGTGGCCACAGTGCACAGTGTGATC 381
 Qy 517 ArgValLeuLeuGlnLysSerArgGluGlnArgPheAspProSerAlaSerLeuAspPhe 536
 Db 382 CGAGTCTGCTGGGCAAGACCCGGGAACAGAGGTTCGACCCCTGCTCTGAGACTTC 441
 Qy 537 LeuTrpAlaCysIleHisValProArgIleTrpGlnGlyArgAspGlnArgTrpProGln 556
 Db 442 CTGTGGGCGCTGATCCATGCTCTCCGATCTGGCAGGGGGGAGACCAAGCCCGCCAG 501
 Qy 557 LysArgArgGluGlnLeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGln 576
 Db 502 AACGGCGGGAGAGCTGGTGGTGGGCTGCGAGGCCCGGAGCTCATCAGCTGTGGAG 561
 Qy 577 LeuIleLeuAlaGlnLagIuThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeu 596
 Db 562 CTGATCTGGCCAGAGCGGAGAGCGGAGACAGGAGGAGACACCCGCTCAGACTC 621
 Qy 597 IleGlnAlaArgLeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArg 616
 Db 622 ATCCAGGCGCGGCTGCCCTGCTGCTCAGCTGCTGCTGGAGACCATGAGATGTCCAG 661
 Qy 617 LysValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArg 636
 Db 682 AAGGTACGAGAGCAGCTGTACAGCTCATCCACAGTGGGAGACAGCGTGGGAGCAG 741
 Qy 637 ArgCysArgAspLeuLeuLeuGlnLeuTyrLeuGlnArgProGluLeuArgValProVal 656
 Db 742 CGTGGCGAGACCTTCTCTGCACTCTTACTTACAGGCGGAGGCTGGGGTCCCGT 801
 Qy 657 ProGluValLeuLeuHisSerGlnGlyAlaAlaSerSerValCysLysLeuAspGly 676
 Db 802 CCGAGGCTCTACTGCACAGCAGAGGGGCTGCCAGCAGCAGCTGTGCAAGCTGAGAG 861
 Qy 677 LeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnAsn 696
 Db 862 CTCTATCCACCGCTTCTATCAGCTCTTGGGAGACACAGGACATCCCGGGCTTGGAGAC 921
 Qy 697 ArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeu 716
 Db 922 CCAAGGGCGGATGCCAGCATGAGCTCCCGGAGCTGGCGGAGCAACCCGCTGCTGCTG 961
 Qy 717 LeuArgHisLeuProMetIleAlaAlaLeuLeuHisGlyArgThrHisLeuAsnPheGln 736
 Db 982 CTGAGGACCTGGCCCATGATCGGGGCTCTCTGACGGCGGACCCACCTCACTTCCAG 1041
 Qy 737 GluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGlnLeuLeuGlnLeu 756
 Db 1042 GAGTTCTGGCAGAGCAACACCTGAGCTCTCTCAGTGTGGGCTGGGAGAGCTG 1101
 Qy 757 LeuGlnProHisValPheArgSerGlnHisGlnGlnAlaLeuTrpAspCysLeuLeuSer 776
 Db 1102 CTGACAGCCGACGCTTCCGAGCGAGCAACAGGGGCGGTGGGAGCTGCTTGTCTC 1161
 Qy 777 PheIleArgLeuLeuLeuAsnArgLysSerSerArgHisLeuAlaPheIleAsn 796
 Db 1162 TATCATCGGCTGCTGTAATTACAGAGAGTCTCCCGCATCTGCTCATCAAC 1221

Qy 797 LysPheValGlnPheIleHisIleTyrIleThrTyrAsnAlaProAlaAlaIleSerPhe 816
 Db 1222 AAGTTTGTGCACTTCAATCCATAGTATCATCTACATGATCCCGAGAGCATCTCTC 1281
 Qy 817 LeuGlnLysHisAlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMet 836
 Db 1282 CTGCAGAGCAGCCCGCCGCTCCAGACCTCTCTCCATCAACAGTGAACCTGGTGTATG 1341.
 Qy 837 LeuLysSerLeuLeuAlaGlyLeuSerLeuProSerArgAspAspArgTrpAspArgGly 856
 Db 1342 CTGAATCTCTCTTGCAGGGCTCAGCTGCTCCAGAGAGGAGACAGACCGAGGCT 1401
 Qy 857 LeuAspGlnGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 876
 Db 1402 CTGAGAGAGAGGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1461
 Qy 877 LeuPheThrProLeuThrAlaAlaGlnMetAlaProTyrMetLysArgLeuSerArgGly 896
 Db 1462 CTGTTCACTCTGACCCGCGGCGGAGATGGCCCTTACATGAACGGCTTCCCGGGG 1521
 Qy 897 GlnThrValGlu-----AspLeuLeuGluValLeu 906
 Db 1522 CAACGGTGGAGGGTGAAGTACGGCTCTTACACCCAGCCAGATCTGCTGGAGTTCTG 1581
 Qy 907 SerAspIleAspGlnMetSerArgArgArgProGluIleLeuSerPhePheSerTrpAsn 926
 Db 1582 AGTACATAGACAGATGATGCTCCGCGAGACCCGAGATCTGAGCTTCTGACCAAC 1641
 Qy 927 LeuGlnArgLeuMetSerSerAlaGlnGlnCysCysArgAsnLeuAlaPheSerLeuAla 946
 Db 1642 CTGACGGGCTGTATGAGCTGCGCGAGAGAGTGTGGCGCAACTGCGCTTACCTGGCC 1701
 Qy 947 LeuArgSerMetGlnAsnSerProSerIleAlaAlaAlaPheLeuProThrPheMetTyr 966
 Db 1702 CTGCGCTCATGAGAAAGACCCCGACATTCAGCGGCTTCTGCGCACGTTCAATGAC 1761
 Qy 967 CysLeuGlnYserGlnAspPheGluValGlnThrAlaLeuArgAsnLeuProGluTyr 986
 Db 1762 TGCCTGGCGCCAGGACCTTGGAGTGTGCACACGCGCTCCGGAACCTGCTGAGTAC 1821
 Qy 987 AlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArgAlaPheLeuValGlyMet 1006
 Db 1822 GCTCTCTGTGGCCAAAGACACCGGCTGTGCTCTCTCACCGGCTTCTGCTGGGCAAG 1881
 Qy 1007 TyrGlyGlnMetAspProSerAlaGlnIleSerGlnAlaLeuArgIleLeuHisMetGln 1026
 Db 1882 TACGGCCAGATGGACCCAGCGGAGATCTCCGAGGCGCTGAGATCCGATATGAG 1941
 Qy 1027 AlaValMet 1029
 Db 1942 GCCGTGATG 1950
 RESULT 6
 AA298062 standard; cDNA; 1559 BP.
 AA298062:
 09-MAY-2000 (first entry)
 Human secreted protein encoding nucleotide sequence SEQ ID NO:56.
 Human: secreted protein; diagnosis; cytostatic; immunosuppressive;
 antiinflammatory; nocotropic; neuroprotective; antiallergic; cancer;
 tumour; neurodegenerative disorder; developmental abnormality; allergy;
 foetal deficiency; blood disorder; immune system disorder; arthritis;
 autoimmune disease; hepatic disease; renal disease; inflammation;
 Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
 infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 reproductive disorder; gastrointestinal disorder; respiratory disorder;
 metabolic disorder; food additive; preservative; ss.

XX Homo sapiens.
 OS
 XX
 PN M0200004140-A1.
 XX
 PD 27-JAN-2000.
 XX
 PF 14-JUL-1999; 99MO-US15849.
 XX
 PR 15-JUL-1998; 98US-00929221.
 PR 15-JUL-1998; 98US-0092922.
 PR 15-JUL-1998; 98US-0092956.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Ehnert R, Olsen HS, Brewer LA, Florence KA, Young PE;
 PI Mucenski M, Endress GA, Soppet DR.
 XX
 DR WPI; 2000-161128/14.
 XX
 PS P-PSDB; AAY87109.
 XX
 PT New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers, neurological or blood disorders
 XX
 PS Claim 1: Page 337-338; 494pp; English.
 XX
 CC The polynucleotide sequences given in AA298017 to AA298108 encode the
 CC human secreted proteins given in AAY87064 to AAY87223. Human secreted
 CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunosuppressive; antiinflammatory; neotropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Human secreted proteins and their polynucleotides can
 CC be used for developing products for the diagnosis or treatment of cancer,
 CC tumours, neurodegenerative disorders, developmental abnormalities and
 CC foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The
 CC proteins or polynucleotides can also be used as food additives or
 CC preservatives. The proteins are also useful for identifying their
 CC binding partners. AA298008 to AA298016 and AAY87063 are sequence used in
 CC the exemplification of the present invention.
 XX
 SQ Sequence 1559 BP; 281 A; 530 C; 478 G; 267 T; 3 other:
 XX
 Alignment Scores:
 Pred. No.: 8,36e-143 Length: 1559
 Score: 2055.00 Matches: 404
 Percent Similarity: 100.008 Conservative: 0
 Best Local Similarity: 100.008 Mismatches: 0
 Query Match: 39.21% Indels: 0
 DB: 21 Gaps: 0
 US-09-929-769-7 (1-1029) x AA298062 (1-1559)
 QY 626 IIEGNGINTTPGIIYASPSERVALLEUGLYARGARGCYSAAGASPLEULEUGINLEU 645
 DB 1 ATTCAGACAGTGGGAGACAGCGTCTGGGCGAGCGCTGCCAGACACTTCTCTGAGCTC 60
 QY 646 TYTLEUGINARGPROGILUENAVAIProValProGluValLeuLeuHisSerGlnIly 665
 DB 61 TACCTAACGGCGCGGAGCTGGGCTGCCCTGAGGTCTTACTACAGCGAAGGG 120
 QY 666 AlAlaserSerServalCysLysLeuaspLysLeuLeuHisArgPheIleThrLeuLeu 685

DB 121 GCTGCACAGACAGCGTCTGCAAGCTGGACGAGCATCATCCAGCTTCATCAGCTCCTT 180
 QY 686 AlAspThrSerAspSerArgAlAlaLeuGluasnArgGlyAlaAspAlaSerMetLacys 705
 DB 181 GCGGACACAGGAGCTCCCGGGCGTTGGAGAACCCAGGGGCGGATGCCAGCATGGCCCTGC 240
 QY 706 ArgLysLeuAlaValAlaHisProLeuLeuLeuArgHisLeuProMetIleAla 725
 DB 241 CGGAAGCTGGGGGCGGCGGACCCGCTGCTGCTCAGGACCTCCCATGATCCCGCGG 300
 QY 726 LeuLeuHisGlyArgThrHisLeuasnPheGlnIupArgGlnGlnAsnHisLeuSer 745
 DB 301 CTCCTGCAGCGGCGGACCCACCTCAACTCCAGGAGTTCCGCGACAGAACCACTGAGCC 360
 QY 746 CysPheLeuHisValLeuGlyLeuLeuGluLeuLeuGlnProHisValPheArgSerGlu 765
 DB 361 TGCTTCCGACAGCTCTGGGCTGCTGGAGCTGCGAGCCGCGAGTGTCCGAGCGAG 420
 QY 766 HisGlnGlyAlaLeuThrAspCysLeuLeuSerPheIleArgLeuLeuasnTyrArg 785
 DB 421 CACCAAGGGGCGGCTGGGAGCTGCTTCTGCTTCATCCGCTGCTGTAATTCAGG 480
 QY 786 LysSerSerArgHisLeuAlaAlaPheIleasnLysPheValGlnPheIleHisLysTyr 805
 DB 481 AAGTCCCTCCGCGCATCTGGCTGCTTCATCAGCAAGTTGTGCAGTTCATCATAGTAC 540
 QY 806 IIEHTTYYRAsnAlaProAlaAlaIleSerPheLeuGlnLysHisAlaAspProLeuHis 825
 DB 541 ATTACTCTCAAAAGCCCGAGCAGCATCTCTCTGCAAGAACAGCGCGAGCCGCTCCAC 600
 QY 826 AspLeuSerPheAspAsnSerAspLeuValMetLysSerLeuLeuAlaGlyLeuSer 845
 DB 601 GACCTGCTCTTGACAAACAGTACTGCTGATGCTGMAATCCCTCTTCAGAGCTCAGC 660
 QY 846 LeuProSerArgAspAspArgThrAspArgGlyLeuaspGlnGlyGlnGlnGlnSer 865
 DB 661 CTGCCACAGCAGGAGCAGACAGCAGCAGCGAGCTGAGCAGAGAGGCGAGAGAGAGAGC 720
 QY 866 SerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThrAlaAlaGlu 885
 DB 721 TCAGCGGCTCTCTCCCTCCCTGTCAGCGTCTCCCTGTCACCCCTCTGACCCGCGCGAG 780
 QY 886 MetAlaProTyrMetLysArgLeuSerArgGlyGlnThrValGlnAspLeuLeuGluVal 905
 DB 781 ATGGCCCCCTACATTAACAGGCTTTCCCGGGCCAAAGCGTGGAGATGCTGAGAGTT 840
 QY 906 LeuSerAspIleaspGluMetSerArgArgArgArgProGluIleLeuSerPhePheSerThr 925
 DB 841 CTGAGTGCATAGACGAGATGTCCCGGCGAGACCCGAGATCTCGAGCTTCTTCGACCC 900
 QY 926 AsnLeuGlnArgLeuMetSerSerArgAlaGlnGlyCysArgAsnLeuAlaPheSerLeu 945
 DB 901 AACCTGCAGCGGCTGATGAGCTCGGCGCGAGGAGTGTTCGCGCAACCTCGCTTCAGCTG 960
 QY 946 AlAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaAlaPheLeuProThrPheMet 965
 DB 961 GCCCTGCGCTCATATCAGAACAGCCCACTTGCAGACCGGCTTCTGCGCCAGCTTCATG 1020
 QY 966 TyrCysLeuGlySerGlnAspPheGlnValValGlnThrAlaLeuArgAsnLeuProGlu 985
 DB 1021 TACTGCTCGGCGACGACGACTTGTAGGTGTGAGAGAGGCGCTCGGAACCTGCTGAG 1080
 QY 986 TyrAlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArgAlaPheLeuValGly 1005
 DB 1081 TACGCTCTCTCTGTGCAAGAGACGCGCTGTGCTGCTCCACCGGCGCTTCTGTGTGGC 1140
 QY 1006 MetTyrGlnGlnMetAspProSerArgGlnIleSerGlnAlaLeuArgIleLeuHisMet 1025
 DB 1141 ATGTACGGCCAGATGAGACCCAGCGCGCAGATCTCGAGGCGCTGAGGATCTGCATATG 1200
 QY 1026 GlnAlaValMet 1029
 DB 1026 GlnAlaValMet 1029

DB	1201	GAGCCGTCGATG	1212
RESULT 7			
AAD11675			
ID	AAD11675	standard; cDNA: 1559 BP.	
XX			
AC	AAD11675;		
XX			
DT	24-SEP-2001	(first entry)	
XX			
DE	Human secreted protein-encoding gene 46	cDNA clone HJBCY35, SEQ ID NO:56.	
XX			
KW	Human; secreted protein; proliferative disorder; cancer; tumour; asthma;		
KW	foetal abnormality; developmental abnormality; haematopoietic disorder;		
KW	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;		
KW	Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;		
KW	psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;		
KW	inflammation; neurological disorder; Alzheimer's disease; food additive;		
KW	angioelastic disorder; kidney disorder; gastrointestinal disorder; allergy;		
KW	cell culture; chemotaxis; vulnery; binding partner identification;		
KW	gene therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	232..1215	
FT		/tag= a	
FT		/product= "Human secreted protein precursor"	
FT	sig_peptide	232..300	
FT		/tag= b	
FT	mat_peptide	301..1212	
FT		/tag= c	
FT		/product= "Mature human secreted protein"	
XX			
PN	WO200151504-A1.		
PD	19-JUL-2001.		
XX			
PR	12-JAN-2001; 2001WO-US00911.		
XX			
PR	13-JAN-2000; 2000US-0482273.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;		
PI	Latieur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;		
PI	Endress GA, Mucenski M, Ebner R;		
XX			
DR	WPI; 2001-425865/45.		
XX	P-PSDB; AAE06086.		
PT	Isolated nucleic acid molecule encoding a human secreted protein is		
PT	used in preventing, treating or ameliorating a medical condition		
XX			
PS	Claim 1; Page 704; 864p; English.		
XX			
CC	AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted		
CC	protein genes; and AAE06041-AAE06132 represent the proteins they encode.		
CC	AAE06133-AAE06205 represent human secreted protein fragments.		
CC	or ameliorating medical conditions, e.g., by protein or gene therapy.		
CC	Pathological conditions can be diagnosed by determining the amount of the		
CC	new protein in a sample or by determining the presence of mutations in		
CC	the new genes. Specific uses are described for each of the 71 genes,		
CC	based on the tissues in which they are most highly expressed, and include		
CC	developing products for the diagnosis or treatment of proliferative		
CC	disorders, cancer, tumours, foetal and developmental abnormalities,		
CC	haematopoietic disorders, diseases of the immune system, AIDS, autoimmune		
CC	neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),		
CC	cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,		
CC	psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,		

QY 866 SerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThrAlaAlaGlu 885
DB 721 TCACCGGCTCTGCGCCCTGCTGAGGCTCTCCCTGTTACACCTCTACCGCGCCGAG 780
QY 886 MetAlaProTyrMetLysArgLeuSerArgGlyGlnThrValGluAspLeuGluVal 905
DB 781 ATGGCCCCCTACATGAAGACGGCTTCCCGGGCCAAACGGTGGAGATCTCTCGAGGTT 840
QY 906 LeuSerAspLeuAspGluMetSerArgArgArgProGluIleLeuSerPhePheSerThr 925
DB 841 CTGAGTACACATGACGAGATGTCCCGGCGAGACCCGAGATCTGAGCTTCTTCGACG 900
QY 926 AsnLeuGlnArgLeuMetSerSerAlaGluGluCysCysArgAsnLeuAlaPheSerLeu 945
DB 901 AACCTGACGGCGCTGATGAGCTCGCGGAGAGTGTTCGCCAACCTCGCTTCAGCCTG 960
QY 946 AlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaPheLeuProThrPheMet 965
DB 961 GCCCTGGCTCCATGACGAGAACGCCCGCATGTGACGCCCTTCTCTCCACGTTCAATG 1020
QY 966 TyrTyrLeuGlySerGlnAspPheGluValValGlnThrAlaLeuArgAsnLeuProGlu 985
DB 1021 TACTGCTGGGACGACGAGCTTGTGAGTGTGACAGAGCGCTTCGGAACCTGCTGAG 1080
QY 986 TyrAlaLeuLeuCysGlnGluHisAlaAlaValLeuLeuHisArgAlaPheLeuValGly 1005
DB 1081 TACGCTCTCCGCTGCCAAGACAGCGGCTGTGCTGCTCCACCGGCTTCTGCTGAGG 1140
QY 1006 MetTyrGlyGlnMetAspProSerAlaGlnIleSerGluAlaLeuArgIleLeuHisMet 1025
DB 1141 ATGACGCGCCAGATGAGACCCGACGCGCAGATCTCCGAGCGCTGAGATCTCGATATG 1200
QY 1026 GluAlaValMet 1029
DB 1201 GAGGCCGTGATG 1212

RESULT 8
ABK69771
ID ABK69771 standard; cDNA; 1559 BP.
AC ABK69771;
XX 15-JUL-2002 (first entry)
DE Human secreted protein gene 46.
XX
KW Human; ss; gene; secreted protein; gene therapy; immunosuppressive;
KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
KW vasculotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
KW virucide; fungicide; ophthalmological; autoimmune disease; neoplasm;
KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; skin aging;
KW epithelial cell proliferation; food additive.
XX
OS Homo sapiens.
XX
PN MO200226931-A2.
XX
PD 04-APR-2002.
XX
PF 24-SEP-2001; 2001MO-US29871.
XX
PR 25-SEP-2000; 2000US-234925P.
XX
PR 12-JAN-2001; 2001MO-US00911.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,
PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,
PI Endress GA, Mucenski M, Ebner R;

XX WPI: 2002-362489/39.
DR P-PSDB; ABG33908, ABG33997.
XX
PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
PT the polypeptides, useful for treating Huntington's disease, sepsis,
PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis,
PT asthma
XX
XX Claim 1; Page 1187-1188; 1478pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (or its
XX fragment, homologue complement or allelic variant) encoding a human
XX secreted protein (and its fragment, domain, epitope, variant, secreted
XX form and species variant). Also included are a recombinant vector
XX comprising the nucleic acid, a recombinant host cell comprising the
XX vector, an antibody against the secreted protein, a recombinant host cell
XX that expresses the secreted protein and a method of identifying a binding
XX partner of the secreted protein. The nucleic acid and protein are used to
XX prevent, diagnose, treat or ameliorate a medical condition in e.g.
XX humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
XX for example autoimmune diseases e.g. rheumatoid arthritis,
XX hyperproliferative disorders e.g. neoplasms of the breast or liver,
XX cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
XX e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
XX Alzheimer's disease, infections caused by bacteria, viruses and fungi and
XX ocular disorders e.g. corneal infection. Many other diseases and
XX disorders are listed in the specification. The polypeptides can also be
XX used to aid wound healing, an epithelial cell proliferation, to prevent
XX skin aging due to sunburn, to maintain organs before transplantation, for
XX supporting cell culture of primary tissues, to regenerate tissues and in
XX chemotaxis. The polypeptides can also be used as a food additive or
XX preservative to increase or decrease storage capabilities. The present
XX sequence encodes a novel human secreted protein of the invention.
XX
S0 Sequence 1559 BP; 281 A; 530 C; 478 G; 267 T; 3 other:
XX
Alignment Scores:
Pred. No.: 8.36e-143 Length: 1559
Score: 2055.00 Matches: 404
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.21% Indels: 0
DB: 24 Gaps: 0
US-09-929-769-7 (1-1029) x ABK69771 (1-1559)
QY 626 IleGlnGlnTyrPglyAspSerValLeuGlyArgArgCysArgAspLeuLeuGlnLeu 645
DB 1 ATCCAGCAAGTGGGAGACAGCTGCTGGGACAGCGCTGCGGAGACTTCTCTGACGCTC 60
QY 646 TyrLeuGlnArgProGluLeuArgValProValProGluValLeuLeuHisSerGlnGly 665
DB 61 TACCTACAGCGGCGGAGCTGGGCTGCCGCTGAGAGTCTCTAGCAGCCAGAGGG 120
QY 666 AlaAlaSerSerSerValCysLysLeuAspGlyLeuIleHisArgPheIleThrLeuLeu 685
DB 121 GCTGCCAGACAGAGCTGTCGAAAGCTGGAGCGGATCATCCAGCGCTATCATCGCTCTT 180
QY 686 AlaAspThrSerAspSerArgAlaLeuGluAsnArgGlyAlaAspAlaSerMetAlaCys 705
DB 181 GCGGACACGAGGAGCTCCCGGCGCTTGGAGAACGAGGCGGATGCGCATGTGCGCTTC 240
QY 706 ArgLysLeuAlaValAlaHisProLeuLeuLeuLeuArgHisLeuProMetIleAlaAla 725
DB 241 CGGAGAGCTGGCGGTGGCGCACCCGCTGCTGCTCAGGACCTGCTCCATGATCCGGCG 300
QY 726 LeuLeuHisGlyArgThrHisLeuAsnPheGlnGluPheArgGlnGlnAsnHisLeuSer 745
DB 301 CTCCTGACAGGCGGACGACCTCAACTTCCAGGAGTTCCGCGACAGAACCACTGAGC 360
QY 746 CysPheLeuHisValLeuGlyLeuLeuGluLeuLeuGlnProHisValPheArgSerGlu 765

OY 959 Alaphelau-ProThrPheMetTyrCysLeuGlySerGlnAspPheGluValValGln 978
 |||||
 DB 277 GCTTTCCTGCGCCNACGTCATCTACTGCTGGGACACCGACTTTGAGGTGTGCAGAC 218
 OY 978 tAlaLeuArgAsnLeuProGluTyrAlaLeuLeuGlyGlnGluHisAlaAlaValLeu 998
 |||||
 DB 217 GGCCCTCGGAGACCTGCGTACGCTCTCTGTGCCAAGACAGCGCGTGTCTGCT 158
 OY 998 uHisArgAlaPheLeuValGlyMetTyrGlyGlnMetAspProSerAlaGlnHisSer 1018
 |||||
 DB 157 CCACGGGCGCTTCGTCGGGCGATGTACGCGCAGATGACCCAGCGCGCATCTCCGA 98
 OY 1018 uAlaLeuArgIleuHisMetGluAlaValMet 1029
 |||||
 DB 97 GGCCCTGAGATGCTGCAATATGAGAGCGCTGATG 64
 RESULT 10
 ABA08362/C
 ID ABA08362 standard; cDNA: 526 BP.
 AC ABA08362;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:138.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; vitrucide; antibacterial;
 KW antifungal; vulnery; antulcer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 PI
 DR WPI: 2001-457740/49.
 DR P-PSDB: ABB11118.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 1: Page 388; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby

CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

SQ Sequence 526 BP; 85 A; 161 C; 181 G; 99 T; 0 other;

Alignment Scores:

Pred. No.:	3 336-57	Length:	526
Score:	891.00	Matches:	174
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	17.00%	Indels:	0
DB:	22	Gaps:	0

US-09-929-769-7 (1-1029) x ABA08362 (1-526)

OY 17 PropioArgAlaAspAspSerGluPheGlnAlaLeuAspIleTrpPheProGluGln 36
 |||||
 DB 524 CCGCTCGAGCCGACAGACGCGAGTTCAGGGCGCTCTGGAATCTGGTTCCGGAGAG 465
 OY 37 LysProLeuProThrAlaPheLeuValAspThrSerGluGluAlaLeuLeuProAsp 56
 |||||
 DB 464 AAGCCACTGCCACCCCTTCTGTGGACACATCGGAGAGCGCTGCTCTCTGAC 405
 OY 57 TrpLeuLysLeuArgMetIleArgSerGluValLeuArgLeuValAspAlaLeuGln 76
 |||||
 DB 404 TGGCTGAGCTGCGCATGTATCCGTTTGAGGTCTCCGCTGGTGAACGCCCTGCGAG 345
 OY 77 AspLeuGluProGlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProValSerSer 96
 |||||
 DB 344 GACCTGAGCCCGACAGACGTGCTGTCTGTGACATGCTTTGGCATCCCGGTGTCAC 285
 OY 97 MetSerLysLeuLeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnTrpLeuGlu 116
 |||||
 DB 284 ATGAGCAAACTCTCCAGTCTTGACGACAGGTGGCCACGACCCCAAGACTCTGGAG 225
 OY 117 GlnAsnIleMetAspLysAsnTyrMetAlaHisLeuValGlnHisGluValArgGly 136
 |||||
 DB 224 CAGAACATATGACACAGAAATTAATGAGCCCATCTGTGTGAGGTCCAGCATGACGGCGC 165
 OY 137 AlaSerGlyGlnTrpPheHisSerLeuLeuThrAlaSerLeuProProArgArgAsp 156
 |||||
 DB 164 GGCTCGGAGGCGACGCTTCCACATCTGTCACAGCCTCCCTGCGCGCCGCGAGAC 105
 OY 157 SerThrGluAlaProLysProLysSerSerProGluGlnProIleGlyGlnValArgIle 176
 |||||
 DB 104 AGCAGAGAGGACCCAAACCAAGACGACCCAGACAGCCCATAGGCCAGGGCGGATT 45

QY 177 ArgValGlyThrGlnLeuArgValLeuGlyProGlnAsp 190
 DB 44 CGGGTGGGACCCAGCTCCGGGTGGGCCCTGAGGACGAC 3
 RESULT 11
 ID ABL05745 standard; cDNA; 6207 BP.
 XX ABL05745;
 AC ABL05745;
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 11717.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001: 2001WO-US09231.
 XX
 PR 23-MAR-2000: 2000US-191637P.
 PR 11-JUL-2000: 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR MPI: 2001-656860/75.
 DR P-PSDB: ABB61642.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 PS Claim 1: SEQ ID NO 11717; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB161675), expressed DNA
 CC sequences (AB101840-AB161675) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 6207 BP; 1572 A; 1715 C; 1543 G; 1377 T; 0 other:
 Alignment Scores:
 Pred. No.: 1.2e-53 Length: 6207
 Score: 863.00 Matches: 288
 Percent Similarity: 43.40% Conservative: 189
 Best Local Similarity: 26.21% Mismatches: 400
 Query Match: 16.47% Indels: 222
 DB: 23 Gaps: 34
 US-09-929-769-7 (1-1029) x ABL05745 (1-6207)
 QY 1 MethHisIleuValValHisAlaMetValIleuLeuThrLeuGlyProProAlaGala 20
 DB 3328 CTCACATTAACATCATTCACGCTTCATATATCTGCTACCTATTC-----AACAGC 3381
 QY 21 AspAspSerGlnGlnAlaLeuLeuAspIleTirPheProGlnGluGluProLeuPro 40
 DB 3382 AACATGCCGAATCGATACCATTTTGGACTATTGGTTCCGCCGAGACGACGAGCC 3441

QY 41 ThrAlaPheLeuValAspThrSerGlnGluAlaLeu---LeuLeuProAspTirPleuLys 59
 DB 3442 GTCCGCTTTTGTGCCAGCATGCGGACAGAGACAGATGCAACTGCTGCCGCTGAGG 3501
 QY 60 LeuArgMetIleArgSerGlnValLeuArgLeuValAspAlaIleuGlnAspLeuGln 79
 DB 3502 CTGAAGATGATCCGCTATCGAGACAGCTATGATGAGGAGCTGCAACAGCTCAACG 3561
 QY 80 ProGlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerLys 99
 DB 3562 CCGGATCAGATCGCTGCTTTGTGCAAACTTTGGCAGCGCGGCTCAACTGATGCTCAAG 3621
 QY 100 LeuLeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGlnAsnIle 119
 DB 3622 CTCTGCAATGCTGCGACCGCTGACTGACAGCACTTGTATGCTGAGAAATGCCATT 3681
 QY 120 MetAspLysAsnTyrMetAlaHisIleValGlnValGlnHisGlnArgGlyAlaSerGly 139
 DB 3682 CTGAACAAAGGCTATGCGCCAGCTATGCAATTCAGACAGCGCGGCTCTTAAGAAAT 3741
 QY 140 GlyGln---ThrPheHisSerLeuLeuThrAlaSerLeuProProArgAspSerThr 158
 DB 3742 GGGCACTACACCGTACAGCGCTTGATCTGCATTCC----- 3777
 QY 159 GluAlaProLysProLysSerSerProGlnGlnProIleGlyGlnGlyArgIleArgVal 178
 DB 3778 -----CACTGCAGAGCTGTGCAGATCTGCC-----AAGATCAGTGTG 3816
 QY 179 GlyThrGlnLeuArgVal-----LeuGlyProGlnAspPheLeuAla 192
 DB 3817 GTATTTCAGAGGCGCTGAATGATGATTACGATTTTCAGACTGCGGATGATAGACC 3876
 QY 193 GlyMetPheLeu-----GlnIlePheProLeuSerProAspProArg 206
 DB 3877 ACTAATCTCTGCGCACCAAGAGATAGCCCAACCATTTCTACGACGCCGAC----- 3930
 QY 207 TyrGlnSerSerProArgProValAlaLeuAlaLeuGlnGlnAlaLeuGlyGlnGlu 226
 DB 3931 ---CAGTTGACTGATGTCGCAAGTACGCTGCCGCTTCAATCAAAACGTGTGGACATG 3987
 QY 227 LeuAlaArgValValGlnGlySerProGlnValProGlyIleThrValArgValLeuGln 246
 DB 3988 CTAGCAAGCCCGAAGCAATAGAGCGGATGTGTACGCCCAATACAGAGGTCTTAGCC 4047
 QY 247 AlaLeuAlaThrLeuLeuSerSerProHisGlyGlyAlaLeuValMetSerMetHisArg 266
 DB 4048 CTG-----GCTTGCAGTGTCAAGATGACCGCTCAC--- 4077
 QY 267 SerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGlnTyrGlnArgCysValPro 286
 DB 4078 -----GCCGACCATTTTAAAG----- 4095
 QY 287 GlnAspThrGlyPheSerSerLeuPheLeuValLeuLeuGlnMetLeuGlnTirPleu 306
 DB 4096 -----ACTTCTTCAGCTCATCTGACAGCGACAGATACATCTGGAG----- 4143
 QY 307 AspSerProGlyValGlnGlyGlyProLeuArgAlaGlnLeuArgMetLeuAlaSerGln 326
 DB 4144 -----AACGCTCTCCAAAGAACTTAGTATGTT----- 4173
 QY 327 AlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeuLeuArgLeuAlaGlnAla 346
 DB 4174 -----AAGCACACATTCGCGACCTCCAGCTGCTGTCGACGAATCCGAA--- 4215
 QY 347 LeuAlaPheArgGlnAspLeuGlnValValSerSerThrValAlaArgAlaIleAlaTir 366
 DB 4216 ---CTATATCATAGAGCTTGTGTTCACTGCTGAGAAACTCTGCGAGATTATGCG--- 4269
 QY 367 LeuArgSerGlyGlnGlnCysSerValGluProAspLeuIleSer-----LysValLeu 384
 DB 4270 -----CAGCAGTTTAAGGGGAATACCGCTTGTGTGCGACGAAGAACGAGATCGTC 4317

QY 385 GlnGlyLeuLeuLeuValArgSerProHisLeuGluLeuLeuThrAlaPhePheSer 404
 Db 4318 CGGGCGATGCTCCAA-----AGCTTCGAT 4341
 QY 405 AlaThrAlaAspAlaAlaSerProPheProAlaCysLysProValValValSerSer 424
 Db 4342 CAGACCAAGAGAGC-----AAGACCGTCCCAAGTCCAGAGC 4380
 QY 425 LeuLeuLeuGlnGluGluProLeuAlaGlyLysProGlyAlaAspGlySer 444
 Db 4381 GACACGCTTCCAC----- 4395
 QY 445 LeuGluAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrpLeuGluLeu 464
 Db 4396 -----AACGGCTCTTCATCGACTGCTGCTCCGAATG 4428
 QY 465 AspProGluValValSerSerCysProAspLeuGlnLeu-----ArgLeuLeuPhe 481
 Db 4429 GATCCCGAGATAGTTCCCT-----CAGCTAATGAAAGAGCGCTTCTGTTC 4476
 QY 482 SerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTrpLeuLeuThr 501
 Db 4477 TCAAG-----TCTCGACGAGATTAGTTCATCTGTCTTC 4515
 QY 502 LeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysLysLeuArgValLeuLeuGly 521
 Db 4516 CTGATCAACCAACCAACCACTGGAGCGACATCGAAGG-----ATTCGCGAGTACCTGTTC 4572
 QY 522 LysSerArgGluGlnArgPheAspProSerAlaSerLeuAspPheLeuThrAlaCysIle 541
 Db 4573 AAGAAATTCATGAAGACTACAGCTACGCCCGTCTCACTACTCTGAGAGGACTGACC 4632
 QY 542 HisValProArgIleLeuGlnArgSerGlnArgGlnArgThrProGlnLysArgArgGlnGlu 561
 Db 4633 ACCAATCCAGACTGTGGAAGAGCGACCGACAGTACATCTCGAAGACGTACGCCGAT 4692
 QY 562 LeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGluLeuIle----- 578
 Db 4693 GCCTTCTTCATGCTGAGGACTTCTGAACTGGAGCGCTTCTCCACTCTTCACGAA 4752
 QY 579 -----LeuAlaGluAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIle 597
 Db 4753 GGGCTTTCGAGGTCAAGTGGACACCAAACTATGATTTCAAGCTGTGCTG----- 4806
 QY 598 GlnAlaArgLeuProLeuLeuLeuSerCysCysGlyAspAspLeuSerValLeuArgLys 617
 Db 4807 -----CGAATGACTACTGTTTAACTCAGACAGAACGACGACCTCATGTGCAAG 4860
 QY 618 ValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArg 637
 Db 4861 GTAATGGACCGCTGAG-----AAAAGTTCGGTTCGCACTACTTG 4902
 QY 638 CysArgAspLeuLeuLeuGlnLeuTrpLeuGlnArgProGluLeuArgValProValPro 657
 Db 4903 AAATGCAAGTCTCCAGCAGATGATCATGTATCCCGCATCAAGTTTCTGAACCG 4962
 QY 658 -----GluValLeuLeuHisSerGlnValAlaAlaSerSerSer 670
 Db 4963 GCGAAACCGCGAGCAGCGCTACAAATTCGCAATCTGAGAGCG----- 5007
 QY 671 ValCysLysLeuAspCysLysLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAsp 690
 Db 5008 -----TGCAGCGCGAGCAAGGTTCACAAATCTGATCAGCTCTGGCGAGTGTGTGGCG 5064
 QY 691 SerArgAlaLeuGluLysAsnArgCysAlaAspAlaSerMetAlaCysArgLysLeuAlaVal 710
 Db 5065 AAAAAGACTTTGAAACCTGTTCACGACGACGACGAGCTGCTCTTGCAGAGCTGCTGCC 5124
 QY 711 AlaHisProLeuLeuLeuLeuArgHisLeuProMetIleAlaAlaLeuLeuHisLysArg 730
 Db 5125 TCCACACCGCTCTCTCTGCGCAGCTGGGGGTCTGTGTCATTTATGCAAGTCTCG 5184
 QY 731 ThrHisLeuAsnPheGlnGluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisVal 750

Db 5185 GCCCAGCTCAGCATGAAGCCCTGCGGAGAGACACACTTCCACGATTTGATACAGATC 5244
 QY 751 LeuGlyLeuLeuGluLeuLeuGlnProHisValPhe-----ArgSerGlnHis 766
 Db 5245 CTGAGAGCGCTGAGGTGTGCAACCACTTTGAGAGAGCGCTACGAAGCGATGC 5304
 QY 767 GlnGlyAlaLeuThrAspCysLeuLeuSerPheIleArgLeuLeuLeuAsnTrpArgLys 786
 Db 5305 CAAAACAGCTG-----TCTGCTACTTCAACTTCTTCAAG-----CACCAAC 5349
 QY 787 SerSerArgHisLeuAlaAlaPheIleAsnLysPheValGlnPheIleHisLysTrpIle 806
 Db 5350 AACGTAAAGAGCGCTGCAAAATGCTGAAACAGTTGTGCAGATGCTCCAAAGCTACATC 5409
 QY 807 ThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHisAlaAspProLeuHisAsp 826
 Db 5410 AACTCAATCCCTCAAGTCTGTCTTCATCGAAGACGTACGTGGCATCTCGAAGGA 5469
 QY 827 LeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeu 846
 Db 5470 CTTCGACCAAGTACACCTCAGTACAGCAAACTCAGATTGTGTCAGGCCGCTGCGCTG 5529
 QY 847 ProSerArgAspArgThrAspArgGlyLeuAspGluGly----- 861
 Db 5530 CTGCACACAAAGTCGACTCGCGCAGCAAGTATGACAGAGAAAGTAAAGTACAGATC 5589
 QY 861 ----- 861
 Db 5590 GATCTGATGATGATTTGATGATTAAGCATCGCGCAGAACCCCTGTGTACAGAGAT 5649
 QY 862 -----GluGluGluSerSerAlaGlySerLeuPro 871
 Db 5650 CCCATGAGATGATTCGGAACACCAACCATCATGATCCAAAGCAGTGGGGTCCCTATG 5709
 QY 872 LeuValSerValSerLeuPheThrProLeuThrAlaAlaGluMetAlaProTrp----- 889
 Db 5710 GTTCTAACCTTGGGTCGTACACGCCGTGACACGACGACATATGCGCGCACTTCTTC 5769
 QY 890 -----MetLysArgLeuSerArgGlyGlnThrValGluAspLeuGluValLeuSer 907
 Db 5770 GATCTGTCAGATCATTAAGACAGTCCAAACGAGACGTGTGTTGGTCCCATGCA 5829
 QY 908 AspIleAspLeuMetSerArgArgArgProGluIleLeuSerPhePheSerThrAsnLeu 927
 Db 5830 GAGCTGAGTGGCTCACTTCCAAAGATATTGTGTATCAACAGAGCTGTTCGAACGATG 5889
 QY 928 GlnArgLeuMetSerSerAlaGluGlyCysArgAsnLeuAlaPheSerLeuAlaLeu 947
 Db 5890 CTCACCTTATATTCGCGCGAGTCCAGATCCGCTCCATGCCCTTCATCATACGATC 5949
 QY 948 ArgSerMetGlnAsnSerPro-----SerIleAlaAlaAlaPheLeuProThrPhe 964
 Db 5950 AGGCATCTGACAGACATCCCGCAATTCGGACATACACTGTGCAACCTTAAGCCTTAC 6009
 QY 965 MetTrpCysLeuGlySerGlnAspPheGluValGlnThrAlaLeuArgAsnLeuPro 984
 Db 6010 ATTCACTGCTGGGGAGACAGACTCGGTGGAGCAAGGACGCGCATTTGACATATGCGCG 6069
 QY 985 GluTrpAlaLeuLeuLeuGlnGlnHisAlaAlaValLeuLeuHisArgAlaPheLeuVal 1004
 Db 6070 GAGATGTCGGTGTCTGGAGGAACAGCAATGACATCTTAACGGTGTGCTGTGCTG 6129
 QY 1005 GlyMetTrpGlyGlnMetAspProSerAlaGlnIleSerGlnAlaLeuArgIleLeu 1023
 Db 6130 GCGTTAAGTCTGCTGGAACACTGGCCACACAGATTGAAGGTTCCTCCAGACTCTA 6186
 RESULT 12
 ABL05744
 ID ABL05744 standard; cDNA; 8355 BP.
 AC XX
 AC ABL05744;
 XX

DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 11714.
 DE Drosophila: developmental biology; cell signalling; insecticide;
 XX Drosophila: gene; ss.
 KM pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 OS
 PN W0200171042-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE.) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PMD, Myers EW;
 PI
 XX WPI: 2001-656860/75.
 DR P-PSDB; AB861641.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Claim 1: SEQ ID NO 11714; 21bp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB857737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 8355 BP; 2141 A; 2150 C; 2051 G; 2013 T; 0 other;

Alignment Scores:
 Pred. No.: 1.81e-53 Length: 8355
 Score: 863.00 Matches: 288
 Percent Similarity: 43.40% Conservative: 189
 Best Local Similarity: 26.21% Mismatches: 400
 Query Match: 16.47% Indels: 222
 DB: 23 Gaps: 34

US-09-929-769-7 (1-1029) x ABL05744 (1-8355)

QY 1 MethisthaleuvalaIhSalameValIleleuLeuThreugLyuProProAla 20
 DB 4476 CTCACATTAACATCATCTGCGCTCATTAATCTGCTCATTTCC-----AACAGC 4529
 QY 21 AspAspserGluPheGlnAlaLeuLeuAspIleTirPheProGluGluLysProLeuPro 40
 DB 4530 AACATGCCGCAATCATGATTTGGACTATTGGTTTCGCCCGGACACACAGCCGCC 4589
 QY 41 ThAlaPheLeuValAspThrSerGluGluAlaLeu---LeuLeuProAspTirPLeuLys 59
 DB 4590 GTGCGCTTCTCTGCCAGCATATGCCAGACAGACAGTGCATCTGCGCCGACGCTGAAG 4649
 QY 60 LeuArgMetIleArgSerGluValLeuArgLeuValAspAlaAlaLeuGlnAspLeuGlu 79
 DB 4650 CTGAAGTGTATCGCTATCGGTGACAGCTAATGAGGACCTTCAACGATCTTAACG 4709
 QY 80 ProGlnIleuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerLys 99
 DB 80 ProGlnIleuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerLys 99

DB 4710 CCGAGTCAGATCGTCTCTTTGTGCAAACTTTGGACGCCGGTCAACTGCATGTCACAG 4769
 QY 100 LeuLeuGlnIlePheLeuAspGlnAlaValAlaIleAspProGlnIleThreugLyuGlnAsnIle 119
 DB 4770 CTCGTGCATATGCTGGACACACCGCTGATCTGAGACAGTTTATCTGGTGAAGATGCCATT 4829
 QY 120 MetAspLysAsnTyrMetAlaIleValIleValGluGlnHisGluArgLysGlyAlaSerGly 139
 DB 4830 CTGACAAAGGCTATCTGCGCCAGCTAATGAGATTCAGACAGCGCGGGTCTTAAGAAAT 4889
 QY 140 GlyGln---ThrPheHisSerLeuLeuThAlaSerLeuProProArgAspSerThr 158
 DB 4890 GGGCAGTCAACGCTACAGCGCTTGATCTGCATTC----- 4925
 QY 159 GluAlaProLysProLysSerSerProGluGlnProIleGlyGlnIleArgLysVal 178
 DB 4926 -----CAGTCGAGACTGTGCCAGATCTGCC-----AAGATCAGTGTG 4964
 QY 179 GlyThrGlnIleuArgVal-----LeuGlyProGluAspLeuAla 192
 DB 4965 GTTATTCAGGAGCGCGTGAATGTGATGATTACGATTTCTCAAGCTCGCATGTAGACCC 5024
 QY 193 GlyMetPheLeu-----GlnIlePheProLeuSerProAspProArg 206
 DB 5025 ACTAAGTCTCTGGCCACCAAGAGGTAGCCCAACATTTCTACAGACGCCGAC----- 5078
 QY 207 TrpGlnSerSerSerProArgProValAlaLeuAlaLeuGlnIleAlaLeuGlyGlnIle 226
 DB 5079 ---CAGTTGACTGAGTCGCGCAAGTGAAGTGCCTGTTGATTCAAAACTGGTGGACATG 5135
 QY 227 LeuAlaArgValAlaGlnIleGlySerProGluValProGlyIleThrValArgValLeuGln 246
 DB 5136 CTGACAAAGCCGCAACGCAATAGCGGATGTGTTAAACCCCTAAGAGAGGCTTAGGCC 5195
 QY 247 AlaLeuAlaThrLeuLeuSerSerProHisGlyAlaLeuValMetSerMetHisArg 266
 DB 5196 GTG-----GTTGACAGTTCACGATGAGCCCTCAC--- 5225
 QY 267 SerHisPheLeuAlaCysProLeuLeuArgGlnLeuGlnIleTyrGlnArgCysValPro 286
 DB 5226 -----CCGTGCATATTTTAAG----- 5243
 QY 287 GlnAspThrGlyPheSerSerLeuPheLeuValLeuLeuGlnIleMetLeuGlnTirPLeu 306
 DB 5244 -----ACTTCTTCAGCTGATGCTGCACAGGACAGATATCATCTGAG----- 5291
 QY 307 AspSerProGlyValGlnIleGlyProLeuArgAlaGlnLeuArgMetLeuAlaSerGln 326
 DB 5292 -----AACGCTCTCCAAAGACCTTAAGTATGTT----- 5321
 QY 327 AlaSerAlaGlyArgArgLeuSerAspValArgGlyLeuLeuArgLeuAlaGluAla 346
 DB 5322 -----AAGCACAATTCGCCGCACTCCAGTCTGCTGCAAAATCCGAA--- 5363
 QY 347 LeuAlaPheArgGlnAspLeuGluValAlaIleSerSerThrValArgAlaValIleAlaThr 366
 DB 5364 ---CTCATATAGAGACTGTGTTCAATGCTGAGAACTCTGCGAGATTATGCG--- 5417
 QY 367 LeuArgSerGlyGlnIleCysSerValGluProAspLeuIleSer-----LysValLeu 384
 DB 5418 -----CAGCAGTTTAAGGGGAATATCCGCTGTGTGCACAGAAACGATCGTC 5465
 QY 385 GlnGlyLeuIleGluValAlaArgSerProHisLeuGlnIleLeuLeuThrAlaPhePheSer 404
 DB 5466 CCGGCAATCTCCCA----- 5489
 QY 405 AlaThrAlaAspAlaAlaSerProPheProAlaCysLysProValAlaValAlaSerSer 424
 DB 5490 CAGACCAAGACAGC-----AAGACGCTGCCCAAGTCCAAAGAC 5528
 QY 425 LeuLeuLeuGlnIleGluGluProLeuAlaGlyGlyLysProGlyLysAspGlyLysSer 444
 DB 5529 GACCAGCTCTTCAC----- 5543

QY 445 LeuGluAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrpLeuLumMetLeu 464
 DB 5544 -----AACGGCTCTTCATGACGTGGCTGTCACAAATG 5576
 QY 465 AspProGluValValSerSerCysProAspLeuLeu-----ArgLeuLeuPhe 481
 DB 5577 GATCCCGAGATGTTCCACT-----CAGCTAATGAAGAGCGGCTTCTGTTC 5624
 QY 482 SerArgArgGlyGlyGlyGlnAlaGlnValProSerPheArgProTyrLeuLeuThr 501
 DB 5625 TCAAG-----TCTGACAGCGAGTTAGTTGTTCTATCTGTGTTCC 5663
 QY 502 LeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeuLeuGly 521
 DB 5664 CTGATCAACACCAACACGACGACGACGACGAAAG-----ATGCGCAGTACTCTTCC 5720
 QY 522 LysSerArgGluGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCysIle 541
 DB 5721 AAGAAATTCATGAGACTACACTACGACCGCTCTCACTACTCTGAGCGACTGACC 5780
 QY 542 HisValProArgIleTrpGlnGlyArgAspGlnArgThrProGlnLysArgArgGlu 561
 DB 5781 ACCAATCCGAACTGTGGAGGACGACGACAAATGTCGAAGAACGTCAGCGCGGAT 5840
 QY 582 LeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGluLeu----- 578
 DB 5841 GCCTTCTTCATGCTGAGGACTTCTGACTGAGCGGCTTCTCCACTTCACTTCCACGA 5900
 QY 579 ---LeuAlaGluAlaGlnThrArgSerGlnAspGlyAspTrpAlaAlaCysSerLeuIle 597
 DB 5901 GGGCTTCCGAGCTGTGGAGGACGACAAACATGATTCACCTGCTGCTCG----- 5954
 QY 598 GlnAlaArgLeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValaArgLys 617
 DB 5955 -----CGATGACCTTACTGTTTAAGCTCACAGAAAGCGACGACCTCAATGCTCAAG 6008
 QY 618 ValThrGlnHisLeuSerGlyCysIleGlnGlnTrpGlyAspSerValLeuGlyArgArg 637
 DB 6009 GTAATGGAGACAGTGAG-----AAAGTTGCGTGTCCGACACTG 6050
 QY 638 CysArgAspLeuLeuLeuGlnLeuTyrLeuGlnArgProGlnLeuArgValProValPro 657
 DB 6051 AAATGTCAGTCTCCAGAGATGTATCATGATCCGCGCATCAAGTTCTGAAACCG 6110
 QY 658 -----GluValLeuHisSerGlnGlyAlaAlaLaserSerSer 670
 DB 6111 GGCAAAACCGGCGAGCGGCTGACAAATTGCAGAATCTGAAGGC----- 6155
 QY 671 ValCysLysLeuAspGlyLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAsp 690
 DB 6156 ---TCCAGGCGGACAGAGGTGTCCAAACATGTATCAGCTGTTGGCGAGTGTGGTGGC 6212
 QY 691 SerArgAlaLeuGlnAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAlaVal 710
 DB 6213 AAAAAGGACTTGGAAACCTGTCCAGCGACCGAGCTGCTCTTCCAGCAAGCTGGTCC 6272
 QY 711 AlaHisProLeuLeuLeuLeuArgHisLeuProMetIleAlaAlaLeuLeuHisGlyArg 730
 DB 6273 TCCACCGCGCTCTCTCTTCCAGCTGGGCGTCTGTCTGCTCATTAATGCAAGGTCCG 6332
 QY 721 ThrHisLeuAsnPheGlnIlePheArgGlnGlnHisHisLeuSerCysPheLeuHisVal 750
 DB 6333 GCCCAGCTCAGCATGAAGCGCTGCGCAGGACGACACACTTCCACGATTTGTACAGATC 6392
 QY 751 LeuGlyLeuLeuGlnLeuGlnProHisValPhe-----ArgSerGlnHis 766
 DB 6393 CTGAGAGAGCTGTGAGCTGTGCAACCACTTTGAGAGAGCGCTACAAAGAACAGATC 6452
 QY 767 GlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeuLeuLeuAsnTrpArgLys 786
 DB 6453 CAAAACACGCTG---TGTGCTACTTCACTTCTTCAAG-----CACACACACG 6497

QY 787 SerSerArgHisLeuAlaAlaPheIleAsnLysPheValGlnPheIleHisLysTyrIle 806
 DB 6498 AACGTAAAGAGCGCTGCCAAATGCTGAACAGATTGTGCAATGCTCCACACCTCAATC 6557
 QY 807 ThrTyrAsnAlaProAlaAlaIleSerPheLeuGlnLysHisAlaAspProLeuHisasp 826
 DB 6558 AACATCAATCCCTCAAGTCTCTGCTCTTCATCGAAGCACTAGCTGGCATCCTGGAAGAA 6617
 QY 827 LeuSerPheAsnAsnSerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeu 846
 DB 6618 CTTCGACCCCACTACACTCTGAGCAAACTGCAAGGTTTGTCACAGGCGCTTCCCTG 6677
 QY 847 ProSerArgAspAspArgThrAspArgGlyLeuAspGluGluGly----- 861
 DB 6678 CTGCAGACAACTGCCACCTCGGCGAGCAAGATTGAGACAGAGAAATTAGTACAGATAC 6737
 QY 861 ----- 861
 DB 6738 GATCTGATGAGCATTTGCGATGTAAGCCATCGCCAGCAGCCGCTGTAACAGAGAT 6797
 QY 862 -----GluGluGlnSerSerAlaGlySerLeuPro 871
 DB 6798 CCCATCGAAGTAATCCGAAACACCCATCGATCCAGACGACAGTAGGCGTCTTATTCG 6857
 QY 872 LeuValSerValSerLeuPheThrProLeuThrAlaAlaGluMetAlaProTyr----- 889
 DB 6858 GTTCAACCTTGCGCTGTACAGCGCGGTGACGACTACAGGACATATTCGCGCATCTTCC 6917
 QY 890 -----MetLysArgLeuSerArgGlyGlnThrValGluAspLeuLeuGluValLeuSer 907
 DB 6918 GATCTGTGCAATGATCATTAAGCGATCCACACGAGGAGAGTGTGTGGTCCCATGACG 6977
 QY 908 AspIleAspGlnMetSerArgArgArgProGluIleLeuSerPhePheSerThrAsnLeu 927
 DB 6978 GAGCTGAGTGGCTCACTTCCAGAGATTTGTTCATCAACGAGCTGTCCAGACACTG 7037
 QY 928 GlnArgLeuMetSerSerAlaGlnGluCysArgAsnLeuAlaPheSerLeuAlaLeu 947
 DB 7038 CTCAACCTTATATTCCTGCGAGGTGCCAGATCCGCTCATGCTCATCTGATGATC 7097
 QY 948 ArgSerMetGlnAsnSerPro-----SerIleAlaAlaPheLeuProThrPhe 964
 DB 7098 AGGCATCTGTAACACAAATCCCGCAATTGCGACATCAACCTGTGCAACCTTAACGCTTAC 7157
 QY 965 MetTyrCysLeuGlySerGlnAspPheGluValAlaGlnThrAlaLeuArgAsnLeuPro 984
 DB 7158 ATTCACTGCTCGGGAGACAGAACTCTGCTGGTGGCAGGACGCGCATTTACAAATCTGCCG 7217
 QY 985 GluTyrAlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArgAlaPheLeuVal 1004
 DB 7218 GAGATGTGGTCTGCTGTCGAGACACAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 7277
 QY 1005 GlyMetTyrGlyGlnMetAspProSerAlaGlnIleSerGluAlaLeuArgIleLeu 1023
 DB 7278 GCGTTGAATGCTGCTGCTGTAACACTGCGCAGCAAGATTAGAAAGTTCTCCAGACTCTA 7334
 RESULT 13
 ID AAL01467 standard; cDNA, 484 BP.
 AC AAL01467;
 DT 21-NOV-2001 (first entry)
 XX Human reproductive system related antigen cDNA SEQ ID NO: 1468.
 DE Human reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy; ss.
 RN Homo sapiens.
 OS Homo sapiens.
 XX WO20015320-A2.
 XX

PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180528.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI, 2001-465570/50.
DR

DR P-PSDB: AAM95497.
XX Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
PS Claim 1; SEQ ID NO 1466; 1297pp + Sequence Listing: English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 484 BP; 90 A; 161 C; 133 G; 84 T; 16 other:
Alignment Scores:
Pred. No.: 4.25e-35 Length: 484
Score: 592.00 Matches: 137
Percent Similarity: 85.09% Conservative: 0
Best Local Similarity: 85.09% Mismatches: 12
Query Match: 11.30% Indels: 14
DB: Gaps: 2
US-09-929-769-7 (1-1029) x AAL01467 (1-484)
QY 812 AIAAIAIIEserPheleuGlnlySHISAlaAspProleuHISAspSerPheaspasn 831
DB 4 GGAGCATCTCTCTCTCTGACAGAGACGCGGCTCCAGACTGCTCTTCCAGAAC 63
QY 832 SeraspLeuValMetleuLysSerleuLeuAlaGlyLeuSerleuProserArqaspas 851
DB 64 AGTGACCTGGTGATCTCTAAATCCCTCTGAGGGCTCAGCCTCCAGCAGGRACGAC 123
QY 852 ArgThrAspArgGlyLeuAspGlnGlyGlnGlyGlnGlyGlnGlyGlnGlyGlnGly 871
DB 124 AGGACCGACCGAGGCTGACAGAGAGGCGAGGAGAGAGCTCAGCGGCTCTTGCCC 183
QY 872 LeuValSerValSerleuPheThrProleuThrAlaAlaGluMetAlaProTyrMetLys 891
DB 184 CTGGTCAGAGCTCTCCCTGTCACCCCTGACCGGCGGCGAGATGGCCCCCTACATGAAA 243
QY 892 ArgLeuSerArgGlyGlnThrValGlu-----Asp 901
DB 244 CGGCTTTCACCGGGCCAAACGGTGGAGGGTGATGACAGCCCTGTTCAMCCAGCCGAT 303
QY 902 LeuLeuGluValLeuSerAspIleAspGlnMetSerArgArgArgProGluIleLeuSer 921
DB 304 CTGCTGGAGGKTCTGAGTACATAGACGATGTCCCGGAGAGACCGAGATCCTGAGAC 363
QY 922 PhePhe-SerThrAsnLeuGlnArgLeuMetSer---SerAlaGlnGluCysArgAs 940
DB 364 TTCTTCTTCGACCAACGTCGACGGCTGATGACTTCGGCCSARNAAGTGTNGCGSAA 423
QY 940 nLeuAlaPheSerleuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaLaph 960
DB 424 CCTTCCTTTCAGCTT-GCCCTGCGSTC-ATGCGAAGAACGCCAGCATTCAGCGCTTT 481
QY 960 e 960
DB 482 C 482
RESULT 14
ABL07340
ID ABL07340 standard; cDNA; 4351 BP.
XX
XX ABL07340;
AC
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 16502.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.

XX
OS Drosophila melanogaster.
XX
XX
PN WO2001/1042-A2.
XX
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US092231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
DR P-PSDB: ABB63237.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
PS Claim 1; SEQ ID NO 16502; 21pp + Sequence Listing: English.
XX
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB57737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 4351 BP; 1235 A; 981 C; 906 G; 1229 T; 0 other:
Alignment Scores:
Pred. No.: 8.32e-07 Length: 4351
Score: 228.00 Matches: 63
Percent Similarity: 45.39% Conservative: 70
Best Local Similarity: 21.50% Mismatches: 114
Query Match: 4.35% Indels: 46
DB: Gaps: 4
US-09-929-769-7 (1-1029) x ABL07340 (1-4351)
QY 773 CysLeuSerPheIleArgLeuLeuAsnTyrArgLysSerSerArgHisLeuAla 792
DB 4 TCCTACTTCACACTTCTTCAG-----CACCAAGACAGTAACAGGCGCTGC 51
QY 793 AlaPheIleAsnLysPheValGlnPheIleHisLysTyrIleThrTyrAsnAlaProAla 812
DB 52 CAATGCTGAGAACAGTTGTGTGTCAGATGCTCCAGGCTCACTCACTCACTCACTCACT 111
QY 813 AlaIleSerPheleuGlnlySHISAlaAspProleuHISAspSerPheaspasnser 832
DB 112 GCTGTGCTCTTATCAGAACAGTACGAGTACCTGAGAGACTTGCACCAAGTACAC 171
QY 833 AspLeuValMetleuLysSerleuLeuAlaGlyLeuSerleuProserArqaspasparg 852
DB 172 TCACTAGGAAACAGTCAAGTTTGTGTCAGGCGGCTTGCCTCTCCAGCAAGTCCGAC 231
QY 853 ThrAspArgGlyLeuAspGlnGly----- 861
DB 232 TCGCGAGAGGAATGAGACGAGAGAGATTAAAGTACAGATGATGATGATGATTC 291
QY 861
DB 292 GATGTAAAGCCATGCGCCAGCAAGCCCTGTGTAACAGAGATCCATGAGTGAATCCG 351

QY 862 -----Gluc1Gluc1SerSerAlaGlySerLeuProLeuValSerLeu 877
 DB 352 CAAMACCCATCGATCCAGACAGACAGTGGCCCTTATCGGTTTACCTTGGGCTG 411
 QY 878 PheThProLeuThraAlaAlaGluMetAlaProTyr-----MetLysArgLeu 893
 DB 412 TTAGCGCGTGCAGACTACACAGGACATATGCGCCGACTTCTCGATCTGGTCAATCAT 471
 QY 894 SerArgGlyGlnThraValGlnAspLeuLeuValLeuSerAspIleAspGluMetSer 913
 DB 472 AAGCACTCCACACAGCAGAGACGTGCTTGGGTCATGAGAGAGTGGAGTGCCTCAGT 511
 QY 914 ArgArgArgProGluIleLeuSerPhePheSerThraAsnLeuGlnArgLeuMetSer 933
 DB 532 TCCAAAGATTTGCTGTTCATCAACAGCTGTTGCAACGACACTTATATTCCTG 591
 QY 934 AlaGluGlyCysArgAsnLeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSer 953
 DB 592 CCGAGTGGCCAGATCCGGGTCATCGCTTCATCATCTGATCGGCACTGGAACCAAT 651
 QY 954 Pro-----SerIleAlaAlaPheLeuProThraPheMetTyrCysLeuGlySer 970
 DB 652 CCCGCAATTCGGACATCAACCTGTGACCCCTTACGCTCAATTCAGTGCCTGCGGAC 711
 QY 971 GlnAspPheGluValValGlnThraAlaLeuArgAsnLeuProGluTyrAlaLeuLeu 990
 DB 712 GAACATCTCTCGGTGGCAGCGAGCCCATTCACAAATCTGCGGAGATGTCGGTCTG 771
 QY 991 GlnGlnIleAlaAlaValLeuLeuLeuIleAspArgAlaPheLeuValGlyMetTyrGlnMet 1010
 DB 772 CAGGAACAGCAATTCACATCTCAACGGTGGCTTCTCGTGGGCTTGAAGTGTGCTG 831
 QY 1011 AspProSerAlaGlnIleSerGluAlaLeuArgIleLeu 1023
 DB 832 AACACTGGCCACCATAGTAAGTTCTCCAGACTCTA 870
 RESULT 15
 AAC11139
 ID AAC11139 standard; cDNA; 119 BP.
 AC AAC11139;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 15214.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 XX gene therapy; chromosome mapping; ss.
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 990S-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1; SEQ ID 15214; 71bp + CD-ROM; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively

CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

CC
 SQ Sequence 119 BP; 19 A; 44 C; 32 G; 24 T; 0 other;

Alignment Scores:

	2.24e-06	Length:	119
Score:	193.00	Matches:	39
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.68%	Indels:	0
DB:	21	Gaps:	0

US-09-929-769-7 (1-1029) x AAC11139 (1-119)

QY 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaPhe 960
 DB 2 CTCGCTTCAGCTGCGCCCTGCGCTCCATGACAGAACAGCCAGCATTCACCCCTTTC 61
 QY 961 LeuProThraPheMetTyrCysLeuGlySerGlnAspPheGluValValGlnThra 979
 DB 62 CTGCCACGTTCTATCTACTGCTGGGACAGCAGACTTTGAGTGTGTCAGACGGCC 118

Search completed: May 13, 2003, 23:42:54
 Job time : 615 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:39:08 ; Search time 47 Seconds
(Without alignments)
4511.117 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 5241
Sequence: 1 MHILVHAMVILLTGPPRA.....MPSAQISEALRIILMEAVM 1029

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	5237	99.9	1377	4	O9P2A8
2	4118	78.6	812	4	O9P2A8
3	3554	67.8	698	4	O8WV40
4	3442.5	65.7	818	11	O91201
5	3244	61.9	640	4	O9Y3W8
6	1816	34.6	673	4	O9BT91
7	1240.5	23.7	310	4	O9NTD1
8	863	16.5	2042	5	O9M1C5
9	713	13.6	162	11	O9CTF7
10	160	3.1	1194	4	O9HTK4
11	156	3.0	2472	16	O91696
12	154	2.9	1638	2	O87001
13	149	2.8	1217	10	O64720
14	145	2.8	1310	4	O9H7L6
15	145	2.8	1346	4	O8TER5
16	139	2.7	891	12	O41442

17	139	2.7	1211	11	O9R0L3	O9R0L3 ratu
18	139	2.7	1235	11	O9R0L5	O9R0L5 ratu
19	139	2.7	1273	11	O9R0L4	O9R0L4 ratu
20	138.5	2.6	1007	4	O8TEE6	O8TEE6 homo
21	135.5	2.6	840	11	O9DC40	O9DC40 mus
22	135.5	2.6	1883	4	O96SC5	O96SC5 homo
23	134.5	2.6	2535	10	O9S240	O9S240 arabi
24	134	2.6	878	12	O91SP3	O91SP3 infect
25	134	2.6	879	12	O8O0X3	O8O0X3 infect
26	133	2.5	990	4	O9BXN3	O9BXN3 infect
27	132.5	2.5	931	4	O96YV6	O96YV6 homo
28	132.5	2.5	934	4	O9P2B0	O9P2B0 homo
29	132.5	2.5	1542	4	O15035	O15035 homo
30	132.5	2.5	1905	4	O96SC6	O96SC6 homo
31	132	2.5	879	12	O9E7P1	O9E7P1 infect
32	132	2.5	2444	5	O9VSK6	O9VSK6 dros
33	131.5	2.5	882	16	O8ZHC7	O8ZHC7 yers
34	131.5	2.5	1386	4	O75064	O75064 homo
35	131	2.5	1302	4	O95155	O95155 homo
36	130	2.5	879	12	O8OHU8	O8OHU8 infect
37	129.5	2.5	840	11	O91VO3	O91VO3 mus
38	129.5	2.5	879	12	O91IX3	O91IX3 infect
39	129.5	2.5	4835	5	O8T5T1	O8T5T1 infect
40	129	2.5	1114	10	O8S6D2	O8S6D2 giza
41	129	2.5	1626	16	O9RY77	O9RY77 delin
42	128.5	2.5	1664	4	O9BZE5	O9BZE5 homo
43	128	2.4	879	12	O8Z630	O8Z630 infect
44	128	2.4	1012	11	O64028	O64028 mus
45	128	2.4	1021	16	O9RV29	O9RV29 delin

ALIGNMENTS

RESULT 1

ID	Q9P2A8	PRELIMINARY:	PRT:	1377 AA.
AC	O9P2A8:			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	KIAA1440 protein (Fragment).			
GN	KIAA1440.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxId=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-BRAIN;			
RC	MEDLINE=20181126; PubMed=10718198;			
RA	Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. XVI.			
RT	The complete sequences of 150 new cDNA clones from brain which code			
RL	for large proteins in vitro."			
DR	EMBL; AB037861; BAA92678.1;			
FT	NON-TER			
SO	SEQUENCE			
Query Match	99.9%; Score 5237; DB 4; Length 1377;			
Best Local Similarity	99.9%; Pred. No. 0;			
Matches 1028; Conservative	0; Mismatches 1; Indels 0; Gaps 0;			
OY	1 MHILVHAMVILLTGPPRADSEFOALIDIFPEPEKPLPTAFVLDTSSEALLLPDWLKL 60			
DB	349 MHILVHAMVILLTGPPRADSEFOALIDIFPEPEKPLPTAFVLDTSSEALLLPDWLKL 408			
OY	61 RHIRESEVRLVDAALODLEPOOLLFFVOSFGIPVSMKSLQFLDOAAVAHQDTLEONIM 120			
DB	409 RHIRESEVRLVDAALODLEPOOLLFFVOSFGIPVSMKSLQFLDOAAVAHQDTLEONIM 468			
OY	121 DKNYMAHVEVQHERGASGGQTFHSLTLTASLPPRRDSTEAPRKSSPEQIGQGRIRVGT 180			


```
RESULT 3
O8WV40 PRELIMINARY: PRT: 698 AA.
AC O8WV40:
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Euteleostomi;
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC018777; AAH18777.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 698 AA; 77709 MW; C917828091BF4778 CRC64;

Query Match 67.8%; Score 3554; DB 4; Length 698;
Best Local Similarity 100.0%; Pred. No. 1.7e-249;
Matches 698; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 332 RLSDVRGGLRLAALAEAFRODLEVVSSITVAVATLRSGECSVEPDLISKYLGILEVR 391
DB 1 RLSDVRGGLRLAALAEAFRODLEVVSSITVAVATLRSGECSVEPDLISKYLGILEVR 60
OY 392 SPHLEELTAFFSATADAAPPACKPVVVSLLQEEEPILAGKRGAGDGSLEAVRLG 451
DB 61 SPHLEELTAFFSATADAAPPACKPVVVSLLQEEEPILAGKRGAGDGSLEAVRLG 120
OY 452 PSSGLLVDMLEMDPEVSSCPDLRLFLSRKKGGAQVPSFRYLLTLFTHGSSWPT 511
DB 121 PSSGLLVDMLEMDPEVSSCPDLRLFLSRKKGGAQVPSFRYLLTLFTHGSSWPT 180
OY 512 LHQCIRVLLGKSREOFDSASIDFLMACIHVPRIWOGDQRTPOKRREELVAVOGPEL 571
DB 181 LHQCIRVLLGKSREOFDSASIDFLMACIHVPRIWOGDQRTPOKRREELVAVOGPEL 240
OY 572 ISLVELLIAEAETRSODGDTAACSLIOARLPULLSCCGDDESVRKYTEHLSCGIOQWD 631
DB 241 ISLVELLIAEAETRSODGDTAACSLIOARLPULLSCCGDDESVRKYTEHLSCGIOQWD 300
OY 632 SVLGRRCRLLIQYLORELPVPEVLLHSEGAASSVCKIDGLIHRFTLLADTSDS 691
DB 301 SVLGRRCRLLIQYLORELPVPEVLLHSEGAASSVCKIDGLIHRFTLLADTSDS 360
OY 692 RALENKGADAMACRKLAAVAHPULLLRHLPMTAALHGHTHLNFQFROONHLSCTLHYL 751
DB 361 RALENKGADAMACRKLAAVAHPULLLRHLPMTAALHGHTHLNFQFROONHLSCTLHYL 420
OY 752 GLELLLOPVPFSEHOGALMDCILSFIRLLNVRKSSRLAAFINKEVFIHKYITYNAP 811
DB 421 GLELLLOPVPFSEHOGALMDCILSFIRLLNVRKSSRLAAFINKEVFIHKYITYNAP 480
OY 812 AATISFLQKADPLHDLSFNSDLVMLKSLAGLSLPSRDDRTDRIHDEGESESSAGSLP 871
DB 481 AATISFLQKADPLHDLSFNSDLVMLKSLAGLSLPSRDDRTDRIHDEGESESSAGSLP 540
OY 872 LVSYSLEPPLTAEMAPYKKRLSRGQTVEDLLEVLSDIDEMSRREPEILSFSTNLQRLM 931
DB 541 LVSYSLEPPLTAEMAPYKKRLSRGQTVEDLLEVLSDIDEMSRREPEILSFSTNLQRLM 600
OY 932 SSAECCRNLAFLSRMNSPISIAAFLPTFMYCLGSQDEEVVOTALRNLPYVALLOQ 991
DB 601 SSAECCRNLAFLSRMNSPISIAAFLPTFMYCLGSQDEEVVOTALRNLPYVALLOQ 660
OY 992 EHAAYLLHRAFLVGMYGOMDPSAQISEALRIHMEAVM 1029
DB 661 EHAAYLLHRAFLVGMYGOMDPSAQISEALRIHMEAVM 1029
```

```
DB 661 EHAAYLLHRAFLVGMYGOMDPSAQISEALRIHMEAVM 698
RESULT 4
O91Z01 PRELIMINARY: PRT: 818 AA.
AC O91Z01:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RA Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC010333; AAH10333.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 818 AA; 90949 MW; A5128C8692BB7663 CRC64;

Query Match 65.7%; Score 3442.5; DB 11; Length 818;
Best Local Similarity 82.3%; Pred. No. 2.7e-241;
Matches 673; Conservative 70; Mismatches 72; Indels 3; Gaps 2;
```

```
OY 215 VALALQALGOELARVYGSPVEVGIITVRYLQALATLLSSPHGALVMSHRSFLACPL 274
DB 1 VALALQALGOELARVYGSPVEVGIITVRYLQALATLLSSPHGALVMSHRSFLACPL 60
OY 275 IROLCOYORCPDPTGESSFLKYLLOMLOMDPVEGGPIRAOQLMLASQASARRLS 334
DB 61 MROLYORVAPDPTGESSFLKYLLOMLOMDPVEGGPIRAOQLMLASQASARRLS 120
OY 335 DVKGGGLRLAALAEAFRODLEVVSSITVAVATLRSGECSVEPDLISKYLGILEVSPH 394
DB 121 DVKGGGLRLAALAEAFRODLEVVSSITVAVATLRSGECSVEPDLISKYLGILEVSPH 180
OY 395 LEEELTAFFSATADAAPPACKPVVVSLLQEEEPILAG-KPKADGSGLEAVRLGSPS 453
DB 181 LEEELTAFFSATADAAPPACKPVVVSLLQEEEPILAG-KPKADGSGLEAVRLGSPS 240
OY 454 SGLLVDMLEMDPEVSSCPDLRLFLSRKKGKGAQVPSFRYLLTLFTHGSSWPT 511
DB 241 SGLLVDMLEMDPEVSSCPDLRLFLSRKKGKGAQVPSFRYLLTLFTHGSSWPT 300
OY 512 LHQCIRVLLGKSREOFDSASIDFLMACIHVPRIWOGDQRTPOKRREELVAVOGPEL 571
DB 301 LHQCIRVLLGKSREOFDSASIDFLMACIHVPRIWOGDQRTPOKRREELVAVOGPEL 360
OY 572 ISLVELLIAEAETRSODGDTAACSLIOARLPULLSCCGDDESVRKYTEHLSCGIOQWD 631
DB 361 ISLVELLIAEAETRSODGDTAACSLIOARLPULLSCCGDDESVRKYTEHLSCGIOQWD 420
OY 632 SVLGRRCRLLIQYLORELPVPEVLLHSEGAASSVCKIDGLIHRFTLLADTSDS 691
DB 421 SVLGRRCRLLIQYLORELPVPEVLLHSEGAASSVCKIDGLIHRFTLLADTSDS 480
OY 692 RALENKGADAMACRKLAAVAHPULLLRHLPMTAALHGHTHLNFQFROONHLSCTLHYL 751
DB 481 RALENKGADAMACRKLAAVAHPULLLRHLPMTAALHGHTHLNFQFROONHLSCTLHYL 540
OY 752 GLELLLOPVPFSEHOGALMDCILSFIRLLNVRKSSRLAAFINKEVFIHKYITYNAP 811
DB 541 GLELLLOPVPFSEHOGALMDCILSFIRLLNVRKSSRLAAFINKEVFIHKYITYNAP 600
OY 812 AATISFLQKADPLHDLSFNSDLVMLKSLAGLSLPSRDDRTDRIHDEGESESSAGSLP 871
DB 601 AATISFLQKADPLHDLSFNSDLVMLKSLAGLSLPSRDDRTDRIHDEGESESSAGSLP 660
OY 872 LVSYSLEPPLTAEMAPYKKRLSRGQTVEDLLEVLSDIDEMSRREPEILSFSTNLQRLM 931
DB 661 LVSYSLEPPLTAEMAPYKKRLSRGQTVEDLLEVLSDIDEMSRREPEILSFSTNLQRLM 931
```

```

Db 661 LVSVLSLPTLVADVAVAPPMKRLSGRAVEDLETLSIDEMSRRRPEVLGFFSTNLQRLM 720
QY 932 SSAECCRNIAFLSLALRSMQNSPSIAAFLPTFMVCLGSDQEVVOTALRNLPEVALLCQ 991
Db 721 SSAECCRNIAFLSLALRSMQNSPSIAAFLPTFMVCLGSDQEVVOTALRNLPEVALLCQ 780
QY 992 EHAVALHRAFLVGMVGMQDPSAQISEALRILMEAVM 1029
Db 781 EHAVALHRAFLVGMVGMQDPSAQISEALRILMEAVM 818

```

RESULT 5

```

Q9Y3W8 ID Q9Y3W8 PRELIMINARY; PRT; 640 AA.
AC Q9Y3W8;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 71.4 kDa protein (Fragment).
GN DKFZ586J0619.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Wambuit R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL050110; CAB43278.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 640 AA: 71375 MW: 1AC887FC2075B21A CRC64;

Query Match 61.9%; Score 3244; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 5e-227;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 397 ELTAFAFSATADAAPPACKPVVVVSSLLQEEPLAGKPGADGSGLEAVRLGSSGL 456
Db 8 ELTAFAFSATADAAPPACKPVVVVSSLLQEEPLAGKPGADGSGLEAVRLGSSGL 67
QY 457 LVDMLEMLDEPVVSSCDLQRLILFSRRKKGQOVPSFRYLLTLTHOSSWPTLQCI 516
Db 68 LVDMLEMLDEPVVSSCDLQRLILFSRRKKGQOVPSFRYLLTLTHOSSWPTLQCI 127
QY 517 RVLLGKREQRFDPASLIDELMACIHVPRIWQGRDQRTPKRREELVLRVQGPPLISLVE 576
Db 128 RVLLGKREQRFDPASLIDELMACIHVPRIWQGRDQRTPKRREELVLRVQGPPLISLVE 187
QY 577 LIIAEATRSQDGDPTAACSLIOARLPLLLSCCGDDESVRKVTGHSGLIOQMGDSVYGR 636
Db 188 LIIAEATRSQDGDPTAACSLIOARLPLLLSCCGDDESVRKVTGHSGLIOQMGDSVYGR 247
QY 637 RCRDILLQTLQRPRLVAVPEVLLHSEGAASSVCKLDGIRHFTLLADTSRRLN 696
Db 248 RCRDILLQTLQRPRLVAVPEVLLHSEGAASSVCKLDGIRHFTLLADTSRRLN 307
QY 697 RGADASMACRLAVAHPLLLRHLPMTAALLHGRTHLNFQEFROQNLSCFLHVLGLLEL 756
Db 308 RGADASMACRLAVAHPLLLRHLPMTAALLHGRTHLNFQEFROQNLSCFLHVLGLLEL 367
QY 757 LQPHVFRSEHOGALMDCLSFIRLLNLRKSSRHAAFINKEVOFIHKYTTYNAPAAISF 816
Db 368 LQPHVFRSEHOGALMDCLSFIRLLNLRKSSRHAAFINKEVOFIHKYTTYNAPAAISF 427
QY 817 LQKHADPLHLSFDNSDLVWLKSLAGLSLPSRDDRTDRDLDEGEESAGSLPLVSVS 876
Db 428 LQKHADPLHLSFDNSDLVWLKSLAGLSLPSRDDRTDRDLDEGEESAGSLPLVSVS 487
QY 877 LFTPLTAENAPYMKRLSRGQTVEDLLEVLSIDEMSRRRPELISFSTNLQRLMSSAE 936

```

```

Db 488 LFTPLTAENAPYMKRLSRGQTVEDLLEVLSIDEMSRRRPELISFSTNLQRLMSSAE 547
QY 937 CCRNLAFSLALRSMQNSPSIAAFLPTFMVCLGSDQEVVOTALRNLPEVALLCQEHAAV 996
Db 548 CCRNLAFSLALRSMQNSPSIAAFLPTFMVCLGSDQEVVOTALRNLPEVALLCQEHAAV 607
QY 997 LLHRAFLVGMVGMQDPSAQISEALRILMEAVM 1029
Db 608 LLHRAFLVGMVGMQDPSAQISEALRILMEAVM 640

```

RESULT 6

```

Q9BT91 ID Q9BT91 PRELIMINARY; PRT; 673 AA.
AC Q9BT91;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 75.1 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC004286; AAH04286.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 673 AA: 75055 MW: 6CE17654F5B3B3DE CRC64;

Query Match 34.6%; Score 1816; DB 4; Length 673;
Best Local Similarity 99.7%; Pred. No. 2.6e-123;
Matches 360; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHILVHAMVILTLGPRADDSFQALDIWPEEKRPPTAFIVDTSEALLIPDWLKL 60
Db 304 MHILVHAMVILTLGPRADDSFQALDIWPEEKRPPTAFIVDTSEALLIPDWLKL 363
QY 61 RMIRSEVRLVDAADOLEFOOLLFFVQSGIPVSSSKLLOFDAVAHDPOTLQGNIM 120
Db 364 RMIRSEVRLVDAADOLEFOOLLFFVQSGIPVSSSKLLOFDAVAHDPOTLQGNIM 423
QY 121 DKNYMAHVEYQHERGASGGQTFHSLTASLPPRRDSTEAPKPKSSPEQPIGGRIRVGT 180
Db 424 DKNYMAHVEYQHERGASGGQTFHSLTASLPPRRDSTEAPKPKSSPEQPIGGRIRVGT 483
QY 181 QLRVLGPEDDLACMFLQIFPLSPDPKQSSSPRVALALQOALGOELARVQGSPEVPGI 240
Db 484 QLRVLGPEDDLACMFLQIFPLSPDPKQSSSPRVALALQOALGOELARVQGSPEVPGI 543
QY 241 TVRYLQALATLLSPHGGALVMSWHRSHFLACPLRLQCYORCVPODTGESSLFLKVL 300
Db 544 TVRYLQALATLLSPHGGALVMSWHRSHFLACPLRLQCYORCVPODTGESSLFLKVL 603
QY 301 QMLQWLDSPEVEGPIRAQRLMLASQASAGRLSDVNGGLRLAELARQDDEVVSSIV 360
Db 604 QMLQWLDSPEVEGPIRAQRLMLASQASAGRLSDVNGGLRLAELARQDDEVVSSIV 663
QY 361 R 361
Db 664 R 664

RESULT 7
Q9NTD1 ID Q9NTD1 PRELIMINARY; PRT; 310 AA.
AC Q9NTD1;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)

```

DE Hypothetical 34.7 kDa protein (Fragment).
 GN DKFZP34C0126.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RA Ansoore W., Wirtner U., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL313735; CAB/0710.1;
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 310 AA; 34728 MW; EDBA23F1D0C08521 CRC64;
 Query Match 23.7%; Score 1240.5; DB 4; Length 310;
 Best Local Similarity 92.6%; Pred. No. 5.2e-82;
 Matches 249; Conservative 2; Mismatches 7; Indels 11; Gaps 1;
 QY 756 LLOPHVSEHOGALMDLSPFIRLLNVRKSSRLAIFINKFQFIHKYITNAPATIS 815
 Db LLOPHVSEHOGALMDLSPFIRLLNVRKSSRLAIFINKFQFIHKYITNAPATIS 60
 QY 816 FLOKHADPLHDLSPDNLDVLMKSLAGLSLPSRDDRTDRCGLDEGEESAGSLPLSV 875
 Db FLOKHADPLHDLSPDNLDVLMKSLAGLSLPSRDDRTDRCGLDEGEESAGSLPLSV 120
 QY 876 SLFPPPLTAENAPYKRLSRGQTVEDLLVSLDIDMSRRPETSIFSTNIQRLMSAE 935
 Db SLFPPPLTAENAPYKRLSRGQTVEDLLVSLDIDMSRRPETSIFSTNIQRLMSAE 180
 QY 936 ECCRRLAASLRSNONSPTIAAFLPTFWCLGSDGFVGTALRNIPVALLCOENAA 995
 Db ECCRRLAASLRSNONSPTIAAFLPTFWCLGSDGFVGTALRNIPVALLCOENAA 240
 QY 996 VLLHRAFLVG-----MYGQMDPS 1013
 Db VLLHRAFLVG-----MYGQMDPS 269
 RESULT 8
 Q9W1C5 PRELIMINARY; PRT; 2042 AA.
 ID Q9W1C5
 AC Q9W1C5
 DF 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 GN CG3173 protein.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Angelidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang O., Chen L.X.,
 RA Brannon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu P., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Bokorova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Paolici J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003462; AAF47145.1;
 DR Flybase; FBgn0034964; CG3173.
 SQ SEQUENCE 2042 AA; 233746 MW; 9144DB7B51F52E55 CRC64;
 Query Match 16.5%; Score 863; DB 5; Length 2042;
 Best Local Similarity 25.4%; Pred. No. 2e-53;
 Matches 280; Conservative 191; Mismatches 402; Indels 230; Gaps 32;
 QY 1 MHLLVHNAVILLTGTGPRPADSEFOALIDIPPEKPLPTFPVDTSEAL-LIPDMK 59
 Db LHINIITHAFILITTS--NSNPESIPILIDYFPFGRRPAPVAFVLSMPQEOYLPMK 1141
 QY 60 LMRISSEVLRLVDAALQDLEPOQLLFVQSGFIPVSSKLLQFLDQAAHDPQTLQONI 119
 Db LMRISSEVLRLVDAALQDLEPOQLLFVQSGFIPVSSKLLQFLDQAAHDPQTLQONI 1201
 QY 120 MDKNMHLVEYQHERGASGQTFHSLTASLPPRRDSTLEAKPKSSPPQPGGRIKRG 179
 Db MDKNMHLVEYQHERGASGQTFHSLTASLPPRRDSTLEAKPKSSPPQPGGRIKRG 1254
 QY 1202 LMKAVLAQLIEIQARGAKNG---HYVQALDLHSHQVTPDLKIS---VVIQAEVL- 1254
 Db LMKAVLAQLIEIQARGAKNG---HYVQALDLHSHQVTPDLKIS---VVIQAEVL- 1254
 QY 180 TQLRVLGPEPDLAGMFLQIFPLSPDRMOSSSPRALAQAOLGELRAVVGSPPEVG 239
 Db TQLRVLGPEPDLAGMFLQIFPLSPDRMOSSSPRALAQAOLGELRAVVGSPPEVG 1255
 QY 240 ITVRVLAQATLILSSPH-----GGALVSMHRSHEFLACPLLRQLQCYOR 283
 Db ITVRVLAQATLILSSPH-----GGALVSMHRSHEFLACPLLRQLQCYOR 1339
 QY 1292 DCRSLIQKLDMLASPNRAVNAITEVLAVGCVTMSRH---ACTFLR----- 1339
 Db DCRSLIQKLDMLASPNRAVNAITEVLAVGCVTMSRH---ACTFLR----- 1339
 QY 284 CVPQDTGFSFLKYLQMLQMLDSDGVGSGPLRAQRLAQAASAGRLSDVRLRL 343
 Db CVPQDTGFSFLKYLQMLQMLDSDGVGSGPLRAQRLAQAASAGRLSDVRLRL 1340
 QY 1340 -----TFSCMLHSDKYHLE-----NALQKMLMF-----KHTPADSSLQK 1377
 Db -----TFSCMLHSDKYHLE-----NALQKMLMF-----KHTPADSSLQK 1377
 QY 344 AEALAFRODLEVVSTVRVAVITLRSGECCSVEPDLS--KVILGLEIERSPHLEELTA 401
 Db AEALAFRODLEVVSTVRVAVITLRSGECCSVEPDLS--KVILGLEIERSPHLEELTA 1378
 QY 402 FFSATADAAPPACKPVVVSLLQEEEPPLAGCGAGDGSLEAVRLGPSSGLLVML 461
 Db FFSATADAAPPACKPVVVSLLQEEEPPLAGCGAGDGSLEAVRLGPSSGLLVML 1419
 QY 462 EMLDEVVSSCPDLQ---RLFSRKKGGQAVPSFRYLLTLTTHOSSWPTLHOCITAV 518
 Db EMLDEVVSSCPDLQ---RLFSRKKGGQAVPSFRYLLTLTTHOSSWPTLHOCITAV 1448
 QY 519 LKGSREGQFDPASDILFACIHPRIWQGRDQRTPPKREELVRYVGPGLISLVELI 578
 Db LKGSREGQFDPASDILFACIHPRIWQGRDQRTPPKREELVRYVGPGLISLVELI 1495
 QY 1496 YLFKNFHEDYDAVYLVNTEALTNPKLMKKGDKYKSNVRRDAPFMILMTSELEPSHT 1555
 Db YLFKNFHEDYDAVYLVNTEALTNPKLMKKGDKYKSNVRRDAPFMILMTSELEPSHT 579
 QY 579 ----LAETRSQDGDGTACSLIQARLPLLLSCCGGDESVKRYTEHLSGCIQMGDSVL 634
 Db ----LAETRSQDGDGTACSLIQARLPLLLSCCGGDESVKRYTEHLSGCIQMGDSVL 634

```

Db 1556 LHEGLSEVKLDSKNYDKLCS-----RMNLLFKLEKRRDLVKVKEHVE-----KSSVS 1605
QY 635 GRRCRDLLDLQYLRPELRVPEVLLHSEGAASSV-----CKRIDGLHREFTLLA 686
Db 1606 DYKLQYVLOQKYNIPRIKFLP-----CKTEQYAKYKQNLKGCQADKYSNNLITCLG 1658
QY 687 DTSDSRALENGDASNAACKRLAVAPLLLRHLPMTAALLHGRTHLNFQEFROONHSC 746
Db 1659 SLVKKPDEFELSTDETLRLKRLASHPLFLRQLGVLSISMOGRAOLSMKRLREHHHR 1718
QY 747 FLHYGLLELLQPHVF-----RSEHGALMDCLLSEIRLLNRYKRSNHLAFIKFQVFI 802
Db 1719 FVQILRLLELQPTIFEBAYKNETQNTL-SCYENFER-----HSSNVEACOMLKKFVOML 1773
QY 803 HKYITVAPAPASIFLQKHADPLHDSFQNSDLYMLKSLAGLSLPSRDDFTDRGDEGC- 861
Db 1774 QAYINYPSSALLFEIQQYGLKELAKYTSGLKQYLVQYVALLQKHSATELDDDEV 1833
QY 862 -----EESSAGSLPLVSVSLFTPLTAEMA 887
Db 1834 KYEYDLDEHDPKPSAKRPVTEDEPIEVNPQTPIDPSSRGPLSVLFLGYSRSNYTDIS 1893
QY 888 PY-----MKRLSRGQTVEDLLEVLSDIDEMSRRPPEILSFSTNLQRLMSSAECCRNIAF 943
Db 1894 PHFLDLVKIIOSTEDVYLCPPMOELCTSKRFVFTNELFERLLNLFSSAQIRSTAF 1953
QY 944 SLALRSQNSP---SIAAFLPTMYCLGSDFEVOTALRNLEPEYALCOEHAVALR 1000
Db 1954 IILIRHLKHNNSDINCLTNINAVIQCLRDENSSVAATAIDNLEMSVLLQEHMIDITV 2013
QY 1001 AFLVGMGQMDPSAQISEALRIL 1023
Db 2014 AFSGLKSCINTGHQIRKVLQTL 2036

RESULT 9
Q9CTF7 AC Q9CTF7 PRELIMINARY; PRT; 162 AA.
ID Q9CTF7 AC Q9CTF7
AC Q9CTF7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 1110015K06Rik protein (Fragment).
GN 1110015K06Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant H.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kiehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Sakurai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welts C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai J.H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).

```

```

DR EMBL: AK003728; BAB22963.1; -
DR MGD; MGI:1915760; 1110015K06Rik.
FT NON_TER 1.
SO SEQUENCE 162 AA; 17939 MW; E343A7CF040CF7A CRC64;
Query Match 13.6%; Score 713; DB 11; Length 162;
Best Local Similarity 87.0%; Pred. No. 4.1e-44;
Matches 141; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
QY 868 GSLPLVSVSLFTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRPPEILSFSTNL 927
Db 1 GSLPLVSVSLFTPLTAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRPPEILSFSTNL 60
QY 928 QRLMSSAECCRNIAFSLALRSQNSPISIAAFLPTMYCLGSDFEVOTALRNLEPEYA 987
Db 61 QRLMSSAECCRNIAFSLALRSQNSPISIAAFLPTMYCLGSDFEVOTALRNLEPEYT 120
QY 988 LLCOEHAVALRHLRAFLVGMGQMDPSAQISEALRILMEAVM 1029
Db 121 LLCOEHAVALRHLRAFLVGMGQMDPSAQISEALRILMEAVM 162

RESULT 10
Q9H7K4 AC Q9H7K4 PRELIMINARY; PRT; 1194 AA.
ID Q9H7K4 AC Q9H7K4
AC Q9H7K4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FLJ00068 protein (Fragment).
GN FLJ00068.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLLEN;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK024475; BAB15765.1; -
DR InterPro: IPR001251; GRAL-TRIO.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000219; RhogEF.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00621; RhogEF; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00325; RhogEF; 1.
DR SMART: SM00516; SEC14; 1.
DR PROSITE: PSS0003; PH_DOMAIN; 1.
FT NON_TER 1.
SO SEQUENCE 1194 AA; 131087 MW; 86Z01A3A0FEACB67 CRC64;
Query Match 3.1%; Score 160; DB 4; Length 1194;
Best Local Similarity 20.9%; Pred. No. 0.01;
Matches 229; Conservative 104; Mismatches 372; Indels 390; Gaps 47;
QY 74 ALDQLEPQQLLFVQSGIPVYSSMSKLLQFLDAVADHPQTLF-QNIMDKN----- 123
Db 58 ATQDEBQ-----GSLPSKRFQLPRADESGDAQRGIVESSVLSSEGRSSVGL 108
QY 124 ---YMAHLVEYHVRGASGGQTFHSLTLASLPERRDSTEAPKPSSEOPIGGRIRVGT 180
Db 109 LCPMSHLSLAQGSDFPGV-----LVGDPGSPKAMPSPGSLPQALSDPVLGS----- 157
QY 181 QLRVLGPEEDLAGFLQITPLSPDKMOSSPPRYVALALQALQGLARVYQSPPEYPT 240
Db 158 ---DPLSESKL-LEAASG-----SGLPKPADCLLQDLQWEL--LASGATLPG-- 202
QY 241 TVRYQLATLTLSSPHGALVMSHRSHPACLP-----LRQLQYQRCV 285

```



```

Db 203 -TRVOGRAVILLCAHSPAMQSECSQELIRLLYLSIRPREVVALGLTVLVARICA 261
QY 286 PODTFSSSL-----FLKVLQMLQMLDSPGVESGLPAOLRMLASQA-----SA 329
Db 262 PSSSLFSGLSQLEAPGAVYVLL-----VSTLKEVPVSGLEQLSOSLTHIPYA 316
QY 330 GRRLLSDVROGLRLAEALAFRODELVSSVTRAVATLRSGEQCSVEPDLISKYLOGLI 388
Db 317 GLPFS-LGGGLPYCHQAMLDFFRRLEALLQNCQAMCA-----LLQGA 358
QY 369 EYRSHLEELLAFSAADADAAPPACKPVVYVSSLLQEEPLAGCPGADGSGLEAV 448
Db 359 E-----SVKAVPQM-----EPG----- 371
QY 449 RLGPSSGLVDMLEMDPEVSSCPDLQLRLFSRRKKGQAQVPSFRYLLTLFTHOSS 508
Db 372 -EVGQLLOQTEVLMQOVDS-----PWLML----- 396
QY 509 WPTLHOCTIRVLKSGREORFDPASLDFLMACIHPRIWQGHORTPOKRREELVRYOG 568
Db 397 -QC-----QGGRELTWLKOEVPVETLSPDYRTAMOKADELYDRVDG 436
QY 569 PELISVLLELLAEATRSODGDTACSLIQARLPLLLSCCGDDSVKRYTHLSGCTIO 628
Db 437 -LHQTLQSLQSMQRIQ-----ALELYQTL-----BARESGLHQIEWMLQ 474
QY 629 WQDSVL-----GRRCRDLLQOLYLQREPLRVPEVLLHSE-----GAASSVCKLDGLIHR 681
Db 475 VQMPRLERAGERSLDMLOAGSFOELVOYAOEYRQSGKFLQPLTGWMAELDPGANG 534
QY 662 ITLLAD-TSDSRALENRGADASMACRIKAVANRLLRLPMIAALLHGRTHLNQEFRO 740
Db 535 LALRAQLTEFSALQOR-----CQRLADARLFLQ-----FREALTVMEGO 576
QY 741 QNHLSCLFVHVLGLLLQPHVRSE-----HOGALM-----DC 773
Db 577 R-----VLALEQERGVVLOQLQHLWTRHPLRPAHRKMMALATGLSEALROEC 628
QY 774 LLSFRLLLNYKRSRHLAFLNKFEVQITNYNAPAIIFLOKHADPLHD-----LS 828
Db 629 RRAMARCDQDTWALQKLEASL-----KLPRVGSSTSLCVSYPARAPRPLRKAYS 680
QY 829 FDNSSLVWKSLAGLSLPSRDRTDGLDEGEBSASGLPLVSVSLFTP----- 880
Db 681 FD-----RNIGOSLSEFACHCHAAITAACRPEAGGALPQASPTVPPGSSDPRSL 733
QY 881 -LTAEMA-----PYMKRLSRGQTVEDLEVLSDIDMSRRRPEILSFSTNIQRL- 930
Db 734 NRIQVLAEWATEREYVRALE--YTMENYFPELLRPDPVQIGRQRAHLFG-NLEKRD 790
QY 931 -MSSAECCRNLAFLSLARSMONSPSTAAFL-----PTFMCLGSQD----- 972
Db 791 FHCHEFLMELEACTR-----HPRRAVAFLRRHRYOFGMYALXSKKPRSDALM 838
QY 973 -FEVOTAL-----RNLPETALLCOEHAVALHRAFLVGMYGOMD 1011
Db 839 SSYGHTEFRKDKQALGDHLDLASYLKPIQRMGKVALLOE-----LARA-----CGGPTQ 889
QY 1012 PSAOISEALRLTME 1026
Db 880 ELSALREASQSLVHQ 904

```

```

RESULT 11
Q91696 PRELIMINARY: PRT: 2472 AA.
AC Q91696:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Still frameshift probable component of chemotactic signal transduction
system.
GN PA0413.

```

```

OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxId=287;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garner L.L., Gollity L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Latbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RT Nature 406:959-964(2000).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
CC EMBL: AE004478; AAG03802.1; -.
DR HSSP: Q56310; 1B30.
DR InterPro: IPR003594; AtPbind_Atpase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR002545; Chem.
DR InterPro: IPR004359; HIS_KIN_sig.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF01584; Chem. 1.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF01627; Hpt; 6.
DR Pfam: PF00072; response_reg; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR ProDom: PD003142; Hpt; 1.
DR SMART: SM00260; Chem; 1.
DR SMART: SM00387; HATPase_C; 1.
DR SMART: SM00073; HPT; 5.
DR SMART: SM00448; REC; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase;
KW complete proteome.
SQ SEQUENCE 2472 AA; 268573 MW; 9886D8A88417A5E CRC64;

```

```

Query Match 3.0%; Score 156; DB 16; Length 2472;
Best Local Similarity 19.1%; Pred. No. 0.056;
Matches 229; Conservative 142; Mismatches 386; Indels 440; Gaps 53;

```

Db 1150 PGSDALLERIALHGERIASIAERYSDAGERRPDLLEAFVGEIMDILDAEDLLERWH 1209
 QY 345 EALAFRODLEVSSTVAATLRS-EGSVDPDLISKVLOGLI----- 388
 Db 1210 EHPDEROEL-----SALREELSTLDGRARHAEIPVEELCAALLALYDAVEEGRILAVSPAF 1265
 QY 389 --EVRSPH-----LEELTA-----FFSATDAASPP 414
 Db 1266 FEEARQAHREALIGMDVAAAGLOVTPPERVAALQELLEAPAAEAVFIPDESIGADDF 1325
 QY 415 -----ACKPVVVVSSLLQEEPLAGKPGADGSL-----BAVRLGPSG-LLV 458
 Db 1326 PEDDEPALPEAVVEEAGAPAEETVPAPAPAGRELBEMVSTFLEAVDILSAGALA 1385
 QY 459 DWLEMDPEVYSSCPDLQRLRLFSRRKKGQAOVPSFRPPLTLFTHQSMPILHOCIRV 518
 Db 1386 QW-QAEPGALSSLSALQ-RDLHTLKGARMAEIAEIGD-----LAHE-----LEA 1428
 QY 519 LIGKSRQRPDPSASLDFLWACHTVPRIMOGROQT-----PQKRREELVLRV---QGP-- 569
 Db 1429 LYEGLVDRRYQHSPQLAGLQACH-DRLAQDLQDSAGQPLPHDLQSIIRPRQGPVA 1487
 QY 570 -----ELISLVELLAEAEETRSDGDTAACSLQARPLPILSCCGDDESVRKVEHLS 623
 Db 1488 EAATPGAESPVETLVAPA---VEEPAAPAAEAEFEEDPELVEFTL---EEGFDILDSAA 1541
 QY 624 GCIOQWGDVY----- 633
 Db 1542 AALQRMWDVDVNTTELEALQDRLTLKGARMAEIGEIGDLAHELEFLEGICGRILRAS 1601
 QY 634 -----LGRRCRDLLQY-----LQR-----PELVPPPEVL-----LH 662
 Db 1602 PALEGLQORCHDELAELVNRGHTLPDGOALIAETIRLRSDPEDELSVPTSLKPLA 1661
 QY 663 SEGASSSVYCKDGLIRFTLLAD-----TSDRALENRGADSMACRLAVA 711
 Db 1662 AKGAADSESLDIFLEADDLLENLELALGRMDGNGDQPLDD----- 1706
 QY 712 HPULLLHRLPMIALHGRHNLNFOEFRQONHLSCLHVLGLLELLOPHVFRSHGALW 771
 Db 1707 --LIRLHTLKGARLAGOTELG-----NLADHLEOHNLDAOQOCAPW 1747
 QY 772 -DCLLSFTLLNLRKSSRLHAFINKEVQFIHKYITYNAPATISFLOKHADPLHDSFD 830
 Db 1748 PDSL-----LDQSGLEGLOVOYDILLERLAE 1775
 QY 831 NSDLVMLKSLAG-----LSLPSRDDTRDTR---GLDEGESEESSAGSLPLVSVSLFTP 880
 Db 1776 DDE-----AGRRPEPAQALVQADD-TDRAVASALAEILRLAPAGA--IMAAEAPR 1824
 QY 881 LTAAEAPYKRLSRGTVEDLELVSDIDEMSRRP-ELISFSTVLQRLMSAE 936
 Db 1825 AAPTTLPEVRKAQEAQ-----EASRRAPQELVKPAELLENLVNLAGE 1870

RESULT 12
 087001 PRELIMINARY: PRT: 1638 AA.
 AC 087001: 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE ChpA.
 GN ChpA.
 OS Pseudomonas aeruginosa.
 OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC WHITCHURCH C.B., Young M.D., Hobbs M., Mattick J.S.,
 RA Pseudomonas aeruginosa chemotactic transduction genes pilL, chpA chpB
 RT

RT and downstream genes chpC, chpD and chpE.
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC - SIMILARITY: TO OTHER PROKARYOTIC SENSOR TRANSDUCTION HISTIDINE
 CC KINASES.
 DR EMBL: U79580: AAC23931.1: -
 DR HSSP: Q56310: 1B30.
 DR InterPro: IPR003594: ATPbind_ATPase.
 DR InterPro: IPR002545: Chew.
 DR InterPro: IPR004359: HIS_KIN_sig.
 DR InterPro: IPR002570: Hpt.
 DR InterPro: IPR001789: Response_reg.
 DR Pfam: PF01584: Chew. 1.
 DR Pfam: PF02518: HATPase_c. 1.
 DR Pfam: PF01627: Hpt. 6.
 DR Pfam: PF00072: response_reg. 1.
 DR ProDom: PD000039: Response_reg. 1.
 DR ProDom: PD003142: Hpt. 1.
 DR SMART: SM00260: Chew. 1.
 DR SMART: SM00387: HATPase_c. 1.
 DR SMART: SM00073: Hpt. 5.
 DR SMART: SM00448: REC. 1.
 KW Kinase; Phosphorylation; Sensory transduction; Transferase.
 SQ SEQUENCE 1638 AA: 178594 MW: 16760818692381E CRC64.

Query Match 2.98; Score 154; DB 2; Length 1638;

Best Local Similarity 19.18; Pred. No. 0.044; Matches 229; Conservative 141; Mismatches 387; Indels 440; Gaps 53;

QY 8 AMVLLTLPGRADSEFEQALDIPPEKPLPAFLVDNSE--EAL--LLPDW----- 57
 Db 12 AONPVVSLPPPADE-----PYDELRVETFEAGEVLETIGRIPLPAKADHD 61
 QY 58 -----LK-LRMIRSEVL-----RLVDAALDLEPQQLLFVQSF 90
 Db 62 REALTEVRRAFHLLKSGRVRALVIGELAMSIENLFNRVLDRIASAEVQ----- 113
 QY 91 GIPVSSMKLLQFLDQAVANDPQLEONINDKNMHALIVEQNHGRGSGGOTFHSLLTAS 150
 Db 114 -----RVVDVALLPELVEE-----FAANQRQRDVDLLAATAHALAKKE 155
 QY 151 ---LPRRDSTEAPKPSSEPOPTG-----OGRLVGTQLRVLGPEDDLAG 193
 Db 156 PLPEPPAPDDGGVP-PEAGADQPSLSDNGVAPRLADPAQAAEAGSDVELLDQ----- 209
 QY 194 MFLQIFPLSPDKMQS-----SSPRVALDQALGQELARVQSPVEYGTIVR 243
 Db 210 -LEIFTEAEETHLEALVGLADCARLPPVTDALQAL-----HTLKSAHMAGITL- 261
 QY 244 VLQALATLLSPHGALVMSHRSHTLA-----CPILRLQCOYORCVPODTGSSFLKY 298
 Db 262 PIAEIAIPLER-----LVKEYSNLAFDLREAEHLHDADQLFRIGLEVOYGAORPLNPI 315
 QY 299 -----LLOMLQWLDSPGV-----EGGPLRAQLBMLASQASAGRRLSDVRCGLRLA 344
 Db 316 PGSDALLERIALHGERIASIAERYSDAGERRPDLLEAFVGEIMDILDAEDLLERWH 375
 QY 345 EALAFRODLEVSSTVAATLRS-EGSVDPDLISKVLOGLI----- 388
 Db 376 EHPDEROEL-----SALREELSTLDGRARHAEIPVEELCAALLALYDAVEEGRILAVSPAF 431
 QY 389 --EVRSPH-----LEELTA-----FFSATDAASPP 414
 Db 432 FEEARQAHREALIGMDVAAAGLOVTPPERVAALQELLEAPAAEAVFIPDESIGADDF 491
 QY 415 -----ACKPVVVVSSLLQEEPLAGKPGADGSL-----BAVRLGPSG-LLV 458
 Db 492 PEDDEPALPEAVVEEAGAPAEETVPAPAPAGRELBEMVSTFLEAVDILSAGALA 551
 QY 459 DWLEMDPEVYSSCPDLQRLRLFSRRKKGQAOVPSFRPPLTLFTHQSMPILHOCIRV 518
 Db 552 QW-QAEPGALSSLSALQ-RDLHTLKGARMAEIAEIGD-----LAHE-----LEA 594
 QY 519 LIGKSRQRPDPSASLDFLWACHTVPRIMOGROQT-----PQKRREELVLRV---QGP-- 569

Db 595 IVEGLVDRHYOHSPQIAGLQACH--DRLAEQDLQSLAGOPLADPHDLQIGIRFRQGPVA 653
 QY 570 -----ELISVELLILAEETRSGDDDTAACSLIQARPLLLSCCGDDESVKRYEHLS 623
 Db 654 EAATPEEASPVVEELVAP--VEEPAAPAAEAEEERDPVELVEFL--EEGFDILDSAA 707
 QY 624 GCIOGWDGSDV----- 633
 Db 708 AALQRMMDVDNITIELEALQORLHTLKGAARMAEIGDGLAHELFYGLCGGRIRAS 767
 QY 634 -----LGRRCRDLLOLY-----LQR---PELRVPEVL---LH 662
 Db 768 PALFGLQORCHDELAEMLAVRGHRTLPGQALIAEIRLRSDPDQLSVPTSVSLKPLA 827
 QY 663 SEGAASSVCKLDGLHRRFTLLAD-----TDSRALENGGADSMCRKLAVA 711
 Db 828 AKGAADSEILDFLEADDDLENELEALGRWDGNGDQPLDD----- 872
 QY 712 HPDLLRLPMIALHGRTHLNFQEFROQNHLSCEFLHYGLLELLQPHVFRSEHOGALN 771
 Db 873 --LLRLHTLTKGARGLAGTELG-----NLAHDEQLHTDAQQCGAPN 913
 QY 772 -DCLSLFTIRLLNRYKSSRHAAFINKEVQFIHKYTYNAPPAISFLQKHADPLHDSFD 830
 Db 914 PDSL-----LDASGLEGLQORQVDLLRERLAE 941
 QY 831 NSDLVWLKSLAG-----LSLPSRDRDR-----GLDEGESESSAGSLPLVSVSLFTP 880
 Db 942 DDE-----AGEPEPAQALVQAD--TDRAVASALAEITRLAPAGA--IMAAEAP 990
 QY 881 LPAEAPMYKRLSRGQVEDLLEVLSDIDEMSRRP--ELISFSTNLQRLMSAAE 936
 Db 991 AAPATLPVPRKQEAQ-----EASRRAPQELVKYPALELMLVNLAGE -1036

RESULT 13

ID 064720 PRELIMINARY; PRT; 1217 AA.
 AC 064720;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE At2g02560 protein.
 GN AT2G02560.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Renning K.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gfoll J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman J.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana."
 RT Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AC004136; AAC18930.1; -;
 SO SEQUENCE 1217 AA; 134621 MW; AA23A6A494EE1EA CRC64;

Query Match 2.8%; Score 149; DB 10; Length 1217;

Best Local Similarity 20.4%; Pred. No. 0.066;
Matches 196; Conservative 137; Mismatches 334; Indels 292; Gaps 49;

QY 46 DTSEBALILPDWIKLRIRREVELRVDAALQDDEPQULLFYVSFGIPVSSMKLLQFLD 105
 Db 409 DTSESS--PKWL--LKEQSVKYSINROLEKSVKTKVGFVSF---LRELVLVLP 457
 QY 106 QAVAHDPQTL-----EONIMKNYMAHL-VEVQHERGASGGQTFSLTASLPDRDSTEA 160
 Db 458 DCLADHIGSLVPGIERLNDKSSISNLKIE-----ALVFVKLYLASHAP----- 501
 QY 161 P-----KPKSSPE-OPIGGRIRV-----GTOLRVLGP----- 187
 Db 502 PVHFHYIKALSSPVLAAGVGRYKVTAEALRVGCELVRYVRPSPAGMDFEKPVHPIYN 561
 QY 188 -----BDDLAGEFLQFPLSPPRMSSSPRPALALQALQELARVY 231
 Db 562 AIMSRLTNODQDEVEKCAITCMGLVISTFGDQIRALPSCLPV--LVDRMGNEITRLT 618
 QY 232 QGSPEVGVIVRVLOALATLLSPFGALVSMHSHFLACPILLRQLOQYORCPDQTF 291
 Db 619 -----AVKAFSVIATSP-----LH-----INLSCVIDHILAEITGF 649
 QY 292 SSLFLKYLQMLQWLQSPGVGGPLRAQLRMLASQASA--GRRL--SDVRGGLRLAEALAF 349
 Db 650 LRRANRVLRQ-----ATLTMNTLVATYGDKIGSEAYEVLVELSLIS 693
 QY 350 RQLEVVSVTVRAVIATLRSGECCSVEPDL--SKVL--QGLIEVRSHLEL--LTA 402
 Db 694 VSDHMTALALE-LCCGLMKGSCSENI SLAVRKVYLPQALTLVKSPLQGLQALLDLOKF 752
 QY 403 FSATA--DAASPFPACKFVVVVSSLL-LQEEPELAGKRP-----GADGG 443
 Db 753 FEALVYHANTSF-----YTLESLSCAKPSQSGVPRKQALYSIACVAVLICLAADKN 807
 QY 444 SLEAVRLGPSSGLVDWLEMLDPEVSSCPDLQRLI---FSRRKG-KCOAQPVSFRPY 498
 Db 808 CSSTVKM-----LMEILKDDSGTNSAKOHLALLSLGEIGRKDLASHAGIET---- 854
 QY 499 LTLFTHOSSMPTLHQCIRVLGKSRQORFDPASLDFEWACIHVPRIMQGRDRTFOKR 558
 Db 855 -IYESFQSPPEETKSAASYALGNIAVGNL--SNYLPFI-----LDQIDNOOK 899
 QY 559 REELVLRVQGPSELISVELLILAEETRSGDDDTAACSLIQARPLLLSCCGDDESVKRY 618
 Db 900 KQYITLH-----SLKEVIVRQ---SVKADQNSVEKILALLNHCESEEGVRNV 948
 QY 619 TEHLSCGIGQWGSVIGRCRDLLQLYLQRPBELRVPEVLLHSEGAASSSVC----- 672
 Db 949 VAECLG-----KMALIEPEKLVAPALQVRTSPAAFTRAVTAVVKY 989
 QY 673 -----KLDGL-----IHRFTTLADSDSRALENRQADASMCRKLAVAH--PLLLRH 719
 Db 990 SYVERPEKIDELIFPQISSFLMLIKD--GDRHV--RRAVASALS---TFHAYKNLLKGL 1042
 QY 720 LPMIALHGRTHLNFQEFROQNHLSCEFLHVL--GLELLQPHVFRSEHOGALWDCILSF 777
 Db 1043 LPELPLLYQGYIAK-KELIRYDLPKFKHYVDGL-----ELKRAAECEVFTL 1090
 QY 778 IRLLYNRKSSRHAAFINKEVQ-----FIHKYTYNAPPAI--SFLQKHADPLH 825
 Db 1091 VDSCLDQVNPSSFIVPLKSGLEDHYDKMLCHILSLADKCSAVLAVLDSLVEPLHK 1150
 QY 826 DLSPD-----NSDLVWLKSLAGLSLPSRDRDRDGLDEGESESSAGSLPL 872
 Db 1151 TISFKPKQADAKOEHRNED--MTRSLRAISSLDRINGVDYSHKFGGLGMKMRKSVPL 1207

RESULT 14

ID 09H7L6 PRELIMINARY; PRT; 1310 AA.
 AC 09H7L6;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE FLJ00056 protein (Fragment).
 GN FLJ00056.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SPLEEN;
 RA Ohara O., Nagase T., Kikuno R., Okumura K.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human
 spleen."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK024463; BAB15753.1;
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000219; RHOGEF.
 DR Pfam: PF00169; PH.1.
 DR SMART: SM00233; PH.1.
 DR SMART: SM00325; RHOGEF.1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 FT NON_TER 1
 SQ SEQUENCE 1310 AA; 141544 MW; E55E30759653C045 CRC64;

Query Match 2.88; Score 145; DB 4; Length 1310;
 Best Local Similarity 19.7%; Pred. No. 0.14;
 Matches 190; Conservative 87; Mismatches 284; Indels 402; Gaps 41;

QY 12 LTTGPPRADSEFQALLDIFPEKPLPTAFIVDT--EALLDPMWKLMIIRSVL 68
 DB 340 LTTTPP-----CPPEPPPSKDTNTLHTLHLSLRD----- 373
 QY 69 RLVDALQDLEPQOLLVFOGFGIVSSMSKILQFLDQ-----VAHDPQLE-- 116
 DB 374 -----LQTLGLSVLLDRQAPRLPALIPALSQLDSDGPLVQRLLILHDLPLELC 427
 QY 117 -----QNTMDKNYM-----AHLYVOHE-----RGASGC--Q 141
 DB 428 GFOGAELVSENDLKRVAKPELQWELGHRDPSHWEIHOEVVRLCRLOQVIGSVRQ 487
 QY 142 TFHSLTASLPRRDSTEARPKSSPEOPIGQGRIVGTQRLVGPEDDLAGMFLQIFPL 201
 DB 488 AIEELGAAEEEEEAAVGMKPLQ-----KVLADRLALQRD--GGALIMRLRS 535
 QY 202 SPDRWSSSPRYVALALQALGQELARVVGSGPEVPGITVRVLAQALATLSSPHGALV 261
 DB 536 TPSSKLEGQGPATLYQEVDEAIHQ--LVR-----LSNLH----- 567
 QY 262 MSMHRSHFLACPLRLQCYORCPQOTGFSSFLKVLQMLQWLDSPGYE-----G 313
 DB 568 -----VQOQEQROQ-----LRLQOVLQWLSGGEQLASFMFG 602
 QY 314 GPLRA-----QLRMLASQASGRRLSDVR----- 337
 DB 603 DTLALQETELRFRAFSAEVOERLAQAREALALEENATSQKVLDFEORLEQVSGILHRA 662
 QY 338 -----GGILRLA-----EALAFRODLEVVSST--VRAVI----- 364
 DB 663 LRLQRFQQAHEWDEGFARLAGAGPREAVLALALRAPEPSAGTFQEKRALALDLGS 722
 QY 365 -ATLRSEQGSVE--POLISKVLOGLIEVRSPHELLTAFSSATADA-----SP 412
 DB 723 PAALREMGRCQARQCELRRIQHLGEEASPR-----GYRRRADGASSGGAQMGPRSP 776
 QY 413 FPACKPVVVVSVLLQEEEPFLAGKPGADGSL-----EAVRLGPS-----GL 457
 DB 777 SPS-----LSSLLPSS-----PGRPARSHCSLAPCEDVEEESPELAPEAGEPPRAVL 827
 QY 458 VDWLEMLDPEVSSCPDLQRLLFSSRRKGQAQAVPSFRPYLLTLFTHQSWMPTLHOCIR 517
 DB 828 IRGLEVTSTEVVD-----RTCSPREH----- 848

QY 518 VLKSRQRPDPSPASDLFMACIHVPRIWQGRDQRTPOKREELVLRQ----- 567
 DB 849 VLLGRAR-----GPDGPMG--VGTPMERKRSISAQRLVSELIAECQDYVATLSEPV 899
 QY 568 -----GPELISVLELLAEFTRSQGDPTACSLIQARL-----PILLS 606
 DB 900 PPPPEL-----TPELKGTMAALASAREELRSFRTHTFLRELQCATHTPL-- 944
 QY 607 CCCDDSESVKRVTEHLSCGICQMDSVLGRRCRDLQLQYLOPELAPVPEVLLHSEGA 666
 DB 945 -----RIGCFILRHGDQ-----SLYQYVYHRIKLEGLALSPSSYGS 984
 QY 667 ASSSVCKLDG-----IHRFTLLADTSDSRALENRGADASMACKLVAHPILL--L 717
 DB 985 MEAGPYLPRALQPLEQTLFRYGRILLE-----LLREAGPELSSBCRALGAAYQLRQEA 1039
 QY 718 RHLPMIALHGRTHINQEFROQNLSCFLHYGLLELQPHYFRSHQALMDCLLSE 777
 DB 1040 RGRDILAVEAVGCEIDKEGQLLHRDPTVIGRKCLR--HVFLEH-----LLLF 1091
 QY 778 IRL 780
 DB 1092 SKL 1094

RESULT 15
 Q8TER5
 ID OTHERS PRELIMINARY; PRT; 1546 AA.
 AC OTHERS;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE FLJ00128 protein (Fragment).
 GN FLJ00128.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SPLEEN;
 RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human
 spleen."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK074057; BAB84883.1;
 FT NON_TER 1
 SQ SEQUENCE 1546 AA; 167672 MW; 7CEFE00757462910 CRC64;

Query Match 2.88; Score 145; DB 4; Length 1546;
 Best Local Similarity 19.7%; Pred. No. 0.18;
 Matches 190; Conservative 87; Mismatches 284; Indels 402; Gaps 41;

QY 12 LTTGPPRADSEFQALLDIFPEKPLPTAFIVDT--EALLDPMWKLMIIRSVL 68
 DB 576 LTTTPP-----CPPEPPPSKDTNTLHTLHLSLRD----- 609
 QY 69 RLVDALQDLEPQOLLVFOGFGIVSSMSKILQFLDQ-----VAHDPQLE-- 116
 DB 610 -----LQTLGLSVLLDRQAPRLPALIPALSQLDSDGPLVQRLLILHDLPLELC 663
 QY 117 -----QNTMDKNYM-----AHLYVOHE-----RGASGC--Q 141
 DB 664 GFOGAELVSENDLKRVAKPELQWELGHRDPSHWEIHOEVVRLCRLOQVIGSVRQ 723
 QY 142 TFHSLTASLPRRDSTEARPKSSPEOPIGQGRIVGTQRLVGPEDDLAGMFLQIFPL 201
 DB 724 AIEELGAAEEEEEAAVGMKPLQ-----KVLADRLALQRD--GGALIMRLRS 771
 QY 202 SPDRWSSSPRYVALALQALGQELARVVGSGPEVPGITVRVLAQALATLSSPHGALV 261
 DB 772 TPSSKLEGQGPATLYQEVDEAIHQ--LVR-----LSNLH----- 803

```
QY 262 MSMSHSHFLACPLRLQICORCVPODGTGSSFLTKVLQMLQMLDPSGVE-----G 313
DB 804 -----VQOQROROC-----LRLQOVLQMLSGPGEOLASFAMPG 838
QY 314 GPLRA-----QLMLASQASAGRRLSDVR----- 337
DB 839 DTLASLOETELFRFSAFVQERLAQAREALALEENATSOQVLDIFEQRLQEVESGLHRA 898
QY 338 -----GGLRLA-----EALAFRODLEVYST---VRAVI----- 364
DB 899 LRLQRFPOQAHFWDEGFARLAGAGPREAVLAALALARRAPEPSAGTFQEMRALALDLS 958
QY 365 -ATLSGECQSYE-PDLISKVLQGLIEVSPHLELLTAFFSATADAA-----SP 412
DB 959 PALREMGRCQARQCELEIRIQHLOEASPR-----GYRRRADGASSGGQWGPSP 1012
QY 413 FPACRNVVVSSLLQEEEPLAGKPGADGSL-----EAVRLGPS-----GLL 457
DB 1013 SPS-----LSSLLPSS---PGPRPAPSHCSLAPCGEDYEPEGPELAPEAGRPRAVL 1063
QY 458 VQMLEMLDEPVVSSCPDLQLRLFLSRKKGGAQVPSFRPYLLTLETHOSSWPTLHQCTR 517
DB 1064 IRLQEVSTEVVD-----RITCSPREH----- 1084
QY 518 VLLGKSREORFDPASLDFLMACIHVPRIWGRDQRTPOKRREELVLRVQ-----SP 567
DB 1085 VLLGRAR-----GPDGPMG-VGTPRMERKRSTISAOQRIVSELACEDYVATLSEPV 1135
QY 568 ---GPELISLVELILAEFTRSODGPTACSLIOARL-----PILLS 606
DB 1136 PPPGPPEL-----TPELRGTWAAALASRERLRSFHRTFLRELQGCATHPJ--- 1180
QY 607 CCCGDESVKVTENHLSGCIQOMGDSVLRRCRDILLQYLQPELRVPVPEVLLHSEGA 666
DB 1181 -----RIGACFLRHGDOF-----SLYAQYVKHKKLENGIALSPSSKGS 1220
QY 667 ASSSVCKLIDGL-----IHRFTLLADTSDSRALENRGADASMACKRLAVHPILL---L 717
DB 1221 MEAGPYLPALQOPLQRLRYGRLEE-----LLREAGPELSEGCRAIGAAVQLLREQEA 1275
QY 718 RHLPMIAALHGRTHLNPFQFRQNHLSGFLVHVLGELLELOPHVFRSEHOGALMDCLSF 777
DB 1276 RGRDLAVEAVRGCEIDLKEGQLHRDPFTVICGRKKCLR-HVFLFEH-----LLLF 1327
QY 778 IRL 780
DB 1328 SKL 1330
```

Search completed: May 7, 2003, 19:42:18
Job time : 62 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 13, 2003, 22:56:01 ; Search time 6720 Seconds
(without alignments) 4456.367 Million cell updates/sec

Title: US-09-929-769-7
Perfect score: 5241
Sequence: 1 MHILVHAWITLLTLPRA.....MDPSAQISEALRIHVEAVM 1029

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O/Cygn2_1/USPRO/spool/US09929769/runat_07052003_161924_8034/app.query.fasta.1.1223
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09929769 @CGN.1.1.4449 @runat.07052003.161924.8034 -NCPU=6 -ICPU=3
-MARK_XLPHY -NO_MMAP -LARGEOURRY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARK.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELPEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pin.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5241	100.0	3437	6	AX468679 Sequence
2	5237	99.9	4434	9	AB037861 Homo sapi
3	4933.5	94.1	4179	9	BC004286 Homo sapi
4	4118	78.6	2769	9	BC013367 Homo sapi
5	3554	67.8	2418	9	BC018777 Homo sapi
6	3442.5	65.7	2730	10	BC010333 Homo sapi
7	3244	61.9	2241	9	HSB800197 Homo sapi
8	3199.5	61.0	14896	2	AC102953 Homo sapi
9	2673.5	51.0	21964	2	AC130221 Mus muscu
10	1359	25.9	1141	9	HSB802043 Homo sapi
11	865	16.5	157851	2	AC020509 Drosophi
12	863	16.5	146432	3	AC004642 Drosophi
13	863	16.5	154840	3	AC009018 Drosophi
14	863	16.5	303823	3	AE003462 Drosophi
15	738.5	14.1	144097	2	AC120290 Rattus no
16	431	8.2	144097	2	AC120290 Rattus no
17	374.5	7.1	6413	3	DM024676 Drosophi
18	186	3.5	22693	3	HSB16153 Homo sapi
19	175	3.3	33896	2	AC116507 Mus muscu
20	171.5	3.3	33896	1	SCK13
21	167	3.2	210017	2	AC102503 Mus muscu
22	164	3.1	4477	9	AK024475 Homo sapi
23	163.5	3.1	14800	6	AX334119 Sequence
24	163.5	3.1	14800	9	HS053204 Human plect
25	163.5	3.1	189050	1	AL646077 Ralstonia
26	162	3.1	14189	9	HSB16153 Homo sapi
27	160.5	3.1	110000	2	LMF16293 Homo sapi
28	160	3.1	90348	1	AF497482 Homo sapi
29	159.5	3.0	13946	1	AE004537 Homo sapi
30	159	3.0	172148	3	LMF16293 Homo sapi
31	157.5	3.0	15580	1	AE004478 Pseudomon
32	157	3.0	33529	6	AR166425 Sequence
33	156.5	3.0	5997	4	AF212149 Oryctolag
34	156.5	3.0	6038	4	AF212148 Oryctolag
35	156.5	3.0	6069	4	AF212147 Oryctolag
36	156.5	3.0	49377	6	188042 Sequence 1
37	156.5	3.0	7425	1	SCU24241 Sorangium c
38	156	3.0	7425	1	AX024365 Sequence
39	156	3.0	7425	6	AX024258 Sequence
40	156	3.0	47713	1	AX024320 Sequence
41	156	3.0	47713	6	AX024213 Sequence
42	155.5	3.0	11904	3	PAU79580 Pseudomonas
43	155	3.0	31252	3	LMF16293 Homo sapi
44	155	3.0	100185	8	AC114899 Ustilago
45	155	3.0	100957	8	AC116560 Ustilago

RESULT 1

ALIGNMENTS

AX468679
LOCUS AX468679 3437 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0216581.
ACCESSION AX468679
VERSION AX468679.1 GI:21901457
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 Gao, W. Q., Polakis, P., Shou, J., Smith, V., Soriano, R., Williams, P. M.,
Wu, T. D. and Zhang, Z.
TITLE Compositions and methods for the diagnosis and treatment of tumor
JOURNAL Patent: WO 0216581-A 3 28-FEB-2002;
Genentech, Inc. (US)
FEATURES
source 1. 3437
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 558 a 1186 c 1098 g 595 t
ORIGIN

Alignment Scores:
Pred. No.: 3,16e-315 Length: 3437
Score: 5241.00 Matches: 1029
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-929-769-7 (1-1029) x AX468679 (1-3437)

QY 1 MethisIleuValValHisAlaMetValIleuLeuThrLeuGlyProProAla 20
DB 46 ATGCACATCTCTGCTGCTGCATGCCATGCTGATCTGATCGGCGCCCTCGAGCC 105
QY 21 AspAspSerLuePheGlnAlaLeuLeuAspIleTyrPheProGluGlyLysProLeuPro 40
DB 106 GAGCAGACGCGAGTTCACAGCGGCTGTCGACATCTGTTCCGAGAGGAGCCACTGCC 165
QY 41 ThAlAPheLeuValAspThrSerGluGluAlaLeuLeuProAspTyrPheLysLeu 60
DB 166 AGCGCTCTCTGCTGCTGCACATCGAGAGCGCTGCTCTCTGCTGCTGCTGCTGCTG 225
QY 61 ArgMetIleArgSerGluValLeuArgLeuValAspAlaAlaLeuGlnAspLeuGluPro 80
DB 226 CGCATATCCGTTCTGAGGTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
QY 81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerLysLeu 100
DB 286 CACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
QY 101 LeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGlnGlnAsnIleMet 120
DB 346 CTCACAGTCTCTGACACAGGAGGCGCCACAGACCCCAAGCTCTGACAGCAGACATCATG 405
QY 121 AspLysAsnTyrMetAlaHisLeuValGluValGlnHisGlnArgGlyAlaSerGlyGly 140
DB 406 GACAAATAATACATGCGCCCTGCTGAGTGCAGCATGAGCGCGCGCTCCGAGGCG 465
QY 141 GlnThrPheHisSerLeuLeuThrAlaSerLeuProProArgArgAspSerThrGlnAla 160
DB 466 CAGACTTTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
QY 161 ProLysProLysSerSerProGlnGlnProIleGlnGlnGlnArgGlyLeaValGlyThr 180
DB 526 CCGAAGCAAGAGCAGCCAGAGCAGCCATGAGCCAGGCGCGGATGCGGCTGGGAGCC 585
QY 181 GlnLeuArgValLeuGlnLysProGlnAspAspLeuAlaGlyMetPheLeuGlnIlePhePro 200
DB 586 CAGCTCCGGGTCTGCTGCGCTGAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645

QY 201 LeuSerProAspProArgTyrPheGlnSerSerSerProArgProValAlaLeuAlaLeuGln 220
DB 646 CTCAGCCCGGAGCCCTCGGTGGCAGAGCTCCAGTCCCGCCCGCGGCTGCTGCTGCTGCTG 705
QY 221 GlnAlaLeuGlnGlnGlnLeuAlaArgValValGlnGlnLysSerProGlnValProGlyIle 240
DB 706 CAGCCCTGGGCGCAGAGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyGlyAlaLeu 260
DB 766 ACGGTGCTGCTGCTGCGAGCGCTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 261 ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGln 280
DB 826 GTGATGTCCATGACACCTTCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLysValLeuLeu 300
DB 886 TACAGCGCTGTGTGCGACAGAGACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
QY 301 GlnMetLeuGlnTyrPheLeuAspSerProGlyValGlnGlyLysProLeuArgAlaGlnLeu 320
DB 946 CAGATGCTGACAGTGGCTGACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeu 340
DB 1006 AGCATGCTTCCAGCCAGCGCTCAGCGCGGCGCAGAGCTCAGTATGCTCAGAGGGGCTC 1065
QY 341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuValValSerSerThrVal 360
DB 1066 CTGCGCTGCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
QY 361 ArgAlaValIleAlaThrLeuArgSerGlyGlnCysSerValGluProAspLeuIle 380
DB 1126 CGTGCCTCATCGCAGCAGCTGAGAGTCTGGGAGAGAGCAGCGAGCGGAGCTGATC 1185
QY 381 SerLysValLeuGlnGlnGlnIleGlnValArgSerProHisLeuGlnGlnLeuLeuThr 400
DB 1186 AGCAAACTCTCCAGAGGGGCTGATCGAGTGAAGTCCCGCCTGAGGAGGAGCTGTGACT 1245
QY 401 AlaPhePheSerAlaThrAlaAspAlaAlaSerProPheProAlaCysLysProValVal 420
DB 1246 GCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1305
QY 421 ValValSerSerLeuLeuLeuGlnGlnGlnGlnProLeuAlaGlyLysProGlyAla 440
DB 1306 GTGGTGAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1365
QY 441 AspGlyGlySerLeuGlnAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTyr 460
DB 1366 GACGGTGGCAGCCTGAGGCGCTGCGGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTG 1425
QY 461 LeuGlnMetLeuAspProGlnValValSerSerCysProAspLeuGlnLeuArgLeuLeu 480
DB 1426 CTGGAATGTGGACCCCGAGGTGTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1485
QY 481 PheSerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTyrLeuLeu 500
DB 1486 TTCTCCCGGAGGAAGGCAAGGTCAAGGCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTG 1545
QY 501 ThrLeuPheThrHisGlnSerSerTyrProThrLeuHisGlnGlnValLeuValLeuLeu 520
DB 1546 ACCCTTTCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1605
QY 521 GlyLysSerArgLysGlnArgPheAspProSerAlaSerLeuAspPheLeuTyrPalAcys 540
DB 1606 GCGAAGAGCCGGGAACAGAGTTCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1665
QY 541 IleHisValProArgIleTyrPheGlnGlyArgAspGlnArgThrProGlnLysArgArgGlu 560
DB 1666 ATCCATGTTCTCTGCAATCTGAGGAGGCGGAGCAGCAGCAGCCCGAGAAAGCGCGGAG 1725
QY 561 GlnLeuValLeuArgValGlnGlnLysProGlnLeuIleSerLeuValGlnLeuIleLeuAla 580


```

|||||
Db 1726 GAGCTGGTCTCCGGGCTCAGGAGCCCGGAGCTCATCAGCTGTGAGCTGATCTCCGACC 1785
Oy 581 GTUAlaGluThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuLeuIleGlnAla 600
Db 1786 GAGCGGAGACCCGGAGAGGAGGACACAGCCGCTGAGCTCTCATCCAGGCGCCG 1845
Oy 601 LeuProLeuLeuLeuSerCysCysGlyAspAspGlySerValArgGlyValIleGln 620
Db 1846 CTGCCCTGCTGCTCAGAGTGCCTGTGGGAGCATGAGAGTGTGAGAGAGTGTGAGAG 1905
Oy 621 HisLeuSerGlyCysIleGlnGlnIleTrpGlyAspSerValLeuGlyArgArgCysArgAsp 640
Db 1906 CACCTGTAGGCTGCATCCAGAGTGGGAGACAGCTGCTGGGAAGCGCTGCGCCAGAG 1965
Oy 641 LeuLeuLeuGlnLeuThrLeuGlnArgProGlyLeuArgValProValProGlyValLeu 660
Db 1966 CTCTCTCTGAGCTCTTACCTACAGCGGCGGAGCTGGGAGTCCGCTGCTGAG3TCTTA 2025
Oy 661 LeuHisSerGlyGlyAlaAlaSerSerValCysLysLeuAspGlyLeuIleHisArg 680
Db 2026 CTGCACAGAGAGGAGGCTGCCAGCAGCAGCAGCTGCGAAGCTGAGCGACTCATCCAGCC 2085
Oy 681 PheIleThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnAsnArgGlyAlaAsp 700
Db 2086 TTCATCAGCGCTCTTCCGAGACACAGCAGCTCCGCGCTGTGAGAACCGAGGCGCGAT 2145
Oy 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuArgHisLeu 720
Db 2146 GCCACATAGGCTCTGCGGAGAGTGGCGGAGCGCACCCGCTGCTGCTGAGCCCTG 2205
Oy 721 PrometIleAlaAlaLeuLeuHisGlyArgThrHisLeuAsnPhedGlnLubPheArgGln 740
Db 2206 CCCATGATCGGGGCTCTCTGCACGCGCGCACCCACACTTCCAGAGATTCGCCGAG 2265
Oy 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGlnProHis 760
Db 2266 CAGAACCCACCTGAGCTGCTCTGCTGACAGCTGCTGAGCTGCTGAGCTGAGCGCGAC 2325
Oy 761 ValPheArgSerGlnHisGlnGlnLeuLeuTrpAspCysLeuLeuSerPheIleArgLeu 780
Db 2326 GTGTTCGAGGAGCAGCAGGAGGCGGCTGTGGAGCTGCTGCTGCTGCTCATCCCGCTG 2385
Oy 781 LeuLeuAsnTrpArgLysSerSerArgHisLeuAlaAlaPheIleAsnLysPheValGln 800
Db 2386 CTGCGAATTTACAGAGAGTCCCTCCGCACTGTGCTGCTTATCAGAGATTTGTGCGAG 2445
Oy 801 PheIleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSerPheLeuGlnLysHis 820
Db 2446 TTCATCCATAGTACATTACTACATGCCCCAGCAGCAGCATCTCTCTGAGAGAGCAC 2505
Oy 821 AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValIleLeuLysSerLeu 840
Db 2506 GCGGAGCCGCTCCAGACCTGCTGCTGCAACAGTGCATGCTGTGAATCCCTC 2565
Oy 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgThrAspArgGlyLeuAspGln 860
Db 2566 CTTCGAGGAGTACGCTGCGCCAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2625
Oy 861 GlyGlnGlnGlnSerSerAlaGlySerLeuProLeuValSerValSerLeuPheThrPro 880
Db 2626 GCGGAGGAGAGAGCTCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2685
Oy 881 LeuThrAlaAlaGlnMetAlaProTrpMetLysArgLeuSerArgGlyGlnThrValGln 900
Db 2686 CTGACCGCGGCGGAGATGAGCCCTTACATGAAAGCGCTTCCCGGAGGAGGAGGAG 2745
Oy 901 AspLeuLeuGlnValLeuSerAspIleAspGlnMetSerArgArgArgProGlnIleLeu 920
Db 2746 GATCGTGTGAGAGTCTGTGAGTACATAGCAGATGTCCCGGCGAGACCGAGATCCG 2805
Oy 921 SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGlnGlnCysArgAsn 940
|||||

```

```

Db 2806 AGCTTCTTCGACCACTGACGCGGCTGATGAGCTGCGGCGGAGAGTGTCCGCAAC 2865
Oy 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerIleAlaAlaIlePhe 960
Db 2866 CTGCGCTTACGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTTTC 2925
Oy 961 LeuProThrPheMetTrpCysLeuGlySerGlnAspPheGlnValAlaGlnThrAlaLeu 980
Db 2926 CTGCCACGCTCATGATGAGCGCTGCGGACAGCAGGACTTTGAGGTGTGACAGAGCCCTC 2985
Oy 981 ArgAsnLeuProGlnTrpAlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArg 1000
Db 2986 CGGAACCTCTCTGAGTACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3045
Oy 1001 AlaPheLeuValGlyMetTrpGlyGlnMetAspProSerAlaGlnIleSerGlnAlaLeu 1020
Db 3046 GCTTCTGCTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3105
Oy 1021 ArgIleLeuHisMetGlnAlaValMet 1029
Db 3106 AGGATCTGCTGATATGAGGCGCGTGATG 3132

RESULT 2
AB037861 4434 bp mRNA linear PRI 14-MAR-2000
LOCUS AB037861
DEFINITION Homo sapiens mRNA for KIAA1440 protein, partial cds.
ACCESSION AB037861
VERSION AB037861.1 GI:7243277
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_11b:pbLuescriptII SK plus
clone:hg02441b.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (sites) Nagase,T., Kikuno,R., Ishikawa,K.I., Hirotsawa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
20181126
JOURNAL MEDLINE
2 (bases 1 to 4434)
REFERENCE Nagase,T., Kikuno,R.
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnaInfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)

FEATURES
source
1..4434
location/Qualifiers
1..4434
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hg02441b"
/tissue_type="brain"
/clone_11b="pbLuescriptII SK plus"
1..4434
/gene="KIAA1440"
1..4134
/gene="KIAA1440"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1440 protein"
/protein_id="BA92678.1"
/db_xref="GI:7243278"
/translaction="AAATSKQITTESSLLLSLTSLDPQGPFRPPPHILIDVKSLSN
QSLRGLHLCGRNDPLIHTIQROASSQSMFWLADLVQSSGLDVLPPVCLCEFL
HDVADHAGSEEDDEGESEKQAKRROQKROLGRLDLDLPKADQETCEVL
FELRLGDSOYAVLAKGSLVLEBSLRGEEKPEPMEDEVDTPVLDGYOMLR
DLPRLEFDSVASTALALQQA IHMETDPORTSAVLITLSOHTPVEROAHSALADV
ARLYVERSTIMSHLFSKSPSASNAVALSALISTSRVYRRRROKSEBEVYSNESQ
DOVFLRWSSGETATMHLIVHAMVITLTLPGRADSEQALDLWPEEKPLPTAF
VDTSEALLLPDWLKLIRIRSEVRLVDAALQDLERPOOLLLEVFQSGIPVSMKTLQ

```

ELDQAVADHPQLEQINIMDKNYMAHLVEVQHERGASGGOTFHSILTAASLPERRDSTEA
 PKKSPDEPQIGQIRVGTOLRVLEPEDLAGMEQIPLSDPDMQSSPPVLA
 LQALGQELARVQSGPVGITRVLAOLATLSPHGLAMHNRHSLACPLLR
 OCOYORVQDPDQSSFLKVLQMLQMLDPSGVGGRLRLMLAQAQASRLS
 DVGRCRLRLAALAPRODLQVYSTRVANTLRSEOCSEVERDLGSKYLOGLIEVRS
 PHUELLITTFESNADADAPFPACKPVVYSSLLDEEERLQKGGAGGSGLEAVL
 GSSGSLQVLMLEKSDREVDSSQDLDLRLFSRKKQGAQVSPRPLTLTETHSS
 WPTLQCLIVLLKLSREORFDPASLDLRLMACIHYPRWQRODPOKRRRELVAV
 QGPBELSLVELLAEARETSODDTMACSLQARLPLLLSCCGDDESVKRYEHLRG
 CIOQMDSVLGRRCRDLLQYLQREPLVPEVYELHSEGAASSGVKLDGLIHERI
 TLLADSDSRALENRGDQSMACRKLAVNHPILLRLHEMIALLRHRTLMNAPFERO
 QNHLSFELVGLIELLOPHRYVRESEHQGLMDCLSFTLLNLRKSSRHNAIFIKF
 VQPIKTYITYNPAISTLQKADPLHDSFSDSDLVMLKSLAGLSLSDRDRTRG
 LDEEERSESSAGSLPLVSVLFTPLAAMAPYMKRLSGQVLEDEVLSDIDEMSR
 RRPELSEFSTNQLORLMSAEBCRCMLAFSLALRSMQNSPSIAAFLPTFMYCLGSOD
 FEVQVOTALNPEYALLCQEHAAVLHRAFLVGMQOMPQSAQISALRLIHEAVY"

BASE COUNT 731 a 1521 c 1419 g 763 t
 ORIGIN

Alignment Scores:
 Pred. No.: 7.91e-315 Length: 4434
 Score: 5237.00 Matches: 1028
 Percent Similarity: 99.90% Conservative: 0
 Best Local Similarity: 99.90% Mismatches: 1
 Query Match: 99.92% Indels: 0
 DB: Gaps: 0

US-09-929-769-7 (1-1029) x AB037861 (1-4434)

QY 1 MethisthaleuValaIhSAlaMeValIleleuLeuthrleuGlyProProArgAla 20
 Db 1045 ATGACATCTCTGCGGTCACATGCCATGCTGTCAGCGCGCCGCGCGAGCC 1104
 QY 21 AspAspSerGluPheGlnAlaLeuLeuAspIleTrpPheProGluGluProLeuPro 40
 Db 1105 GACGACAGCAGTCCAGCGCGCTGTCGACATCTGCTCCGAGGAGAACCCCTGCC 1164
 QY 41 ThrAlaPheLeuValAspThrSerGluGluAlaLeuLeuProAspTrpLeuLeu 60
 Db 1165 ACCGCTCTCTGCGGACACATCGAGAGCGCGCTGCTCTGCTGCTGCTGCTGCTG 1224
 QY 61 ArgMetIleArgSerGluValIleuArgLeuValAspAlaAlaLeuGlnAspLeuGluPro 80
 Db 1225 CGCATATCTGCTGTAGAGTCTCCGCTGTCGAGCCGCCCTGCGAGACCTGAGCCG 1284
 QY 81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerLeu 100
 Db 1285 CACGACGCTCTGCTGTGTCAGCTGCTGTGGCATCCCGGTGTCACATGAGCAACTG 1344
 QY 101 LeuGlnPheLeuAspGlnAlaValAlaIhSAlaAspProGlnThrLeuGlnGlnAspIleMet 120
 Db 1345 CTCACATCTCTGCGGACGAGTGGCCGACGCCCTGCGAGCACTGCGAGAACATCATG 1404
 QY 121 AspLysAsnTrpMetAlaIhSLeuValGlnIhSAlaArgGlnValAspSerGlyGly 140
 Db 1405 GACAAGATTCATGAGCCGACCTGCTGAGGTCACGATAGCGCGCGCTCCGGAGGC 1464
 QY 141 GlnThrPheIhSerLeuLeuThrAlaSerLeuProProArgArgAspSerThrGluAla 160
 Db 1465 CAGACTTTCACACTCTGTGTCACAGCTCCCTGCGCGCGCGAGACAGACAGAGCA 1524
 QY 161 ProLysProLysSerSerProGluGlnProIleGlyGlnGlyArgIleArgValGlyThr 180
 Db 1525 CCCAAACCAAGAGACCCAGAGCAGCCATAGGCGGCGGATTCGGGTGGGAGCC 1584
 QY 181 GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGlnIlePhePro 200
 Db 1585 CAGCTCGGGGTGCTGGGCGCTGAGAGACCTGGCTGGCATGTTCCTCCGAGATTTTCCG 1644
 QY 201 LeuSerProAspArgTrpGlnSerSerSerProArgProValAlaLeuGln 220
 Db 1645 CTGAGCCCGGAGCCCTGCTGGCAGAGCTCCAGTCCCGCCCGCTGCGCTCCGAG 1704

QY 221 GlnAlaLeuGlnGlnGlnLeuAlaArgValValGlnGlySerProGluValProGlyIle 240
 Db 1705 CAGGCGCTGGGCGAGAGCTGGCGCGGCTGCTCAGAGGAGCCCGAGGTGCCGGCATC 1764
 QY 241 ThrValArgValLeuGlnAlaLeuAlaThrLeuLeuSerSerProIhSAlaLeu 260
 Db 1765 ACCGTGGTCTCTGCGAGGCGCTGCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824
 QY 261 ValMetSerMetIhSArgSerIhSPhleuAlaCysProLeuLeuArgGlnIleCysGln 280
 Db 1825 GTATGTCATTCAGCAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1884
 QY 281 TyrGlnArgCysValProGlnAspThrGlyPheSerSerLeuPheLeuValLeuLeu 300
 Db 1885 TACCAGCGCTGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1944
 QY 301 GlnMetLeuGlnTrpLeuAspSerProGluValGlnGlyLeuProLeuArgAlaGlnLeu 320
 Db 1945 CAGATGCTGAGTGGCTGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2004
 QY 321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeu 340
 Db 2005 AGATGCTTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2064
 QY 341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGlnValValSerSerThrVal 360
 Db 2065 CTGGCGCTGGCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2124
 QY 361 ArgAlaValIleAlaThrLeuArgSerGlyGlnGlnCysSerValGluProAspLeuIle 380
 Db 2125 CGTCCGCTCATCCGACCTGAGGTCTGGGAGCAGTGCAGCTGGAGCGCGAGCTGATC 2184
 QY 381 SerLysValLeuGlnGlnGlyLeuIleGluValArgSerProIhSLeuGlnGluLeuLeuThr 400
 Db 2185 AGCAAACTCTCCAGCGGCTGATGAGTGAAGTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2244
 QY 401 AlaPheSerSerAlaThrAlaAspAlaAlaSerProPheProAlaCysLysProValVal 420
 Db 2245 GCATCTTCTCTGCGACCTGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2304
 QY 421 ValValSerSerLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 440
 Db 2305 GTGCTGAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2364
 QY 441 AspGlyGlySerLeuGlnAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrp 460
 Db 2365 GACGCTGCGACCTGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2424
 QY 461 LeuGlnMetLeuAspProGluValValSerSerCysProAspLeuGlnLeuArgLeuLeu 480
 Db 2425 CTGGAATAAGCTGGACCCCGAGGTGTCAGCAGCTGCCCCGACCTGCGAGCTCAGGCTGCTGCTC 2484
 QY 481 PheSerArgArgLysGlyLysGlyGlnAlaGlnValProSerPheArgProTrpLeuLeu 500
 Db 2485 TTCTCCCGGAGGAAGGCAAGGTCAGAGCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2544
 QY 501 ThrLeuPheThrIhSInSerSerTrpProThrLeuIhSInCysIleArgValLeuLeu 520
 Db 2545 ACCCTTTCACGATCAGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2604
 QY 521 GlyLysSerArgGlnGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys 540
 Db 2605 GGCAGAGCGCGGGAACAGAGATTCGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2664
 QY 541 IleIhSValProArgIleTrpGlnGlyArgAspGlnArgThrProGlnLysArgArgGlu 560
 Db 2665 ATCCATTTCTCTGCACTCTGCGAGGCGCGGAGCAGCGCAGCCGCGAGAGCGGCGGAG 2724
 QY 561 GlnLeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGlnLeuIleLeuAla 580
 Db 2725 GAGCTGCTGCTGGGCGTCCAGAGGCGCGGAGCTCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2784
 QY 581 GlnAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg 600

```

|||||
Db 2785 GAGGGGAGACCGGAGCGAGGAGCGGAGCGCCCTGAGCCCTCATCCAGGACCGG 2844
Oy 601 LeuProLeuLeuLeuSerCysCysCysGlyAspAspGlnSerValArgLysValPheGlu 620
Db 2845 CTGCCCCCTGCTCTGAGCTGCTGCTGGGAGATGATGATGATGATGATGATGATGATGAT 2904
Oy 621 HisLeuSerGlyCysL1eGlnGlnInTPGlyAspSerValLeuGlyArgGlyCysAsp 640
Db 2905 CACCTGTGAGGCTGCATCCAGAGTGGGAGAGACAGCTGCTGGGAGGCGCTGCCAGAC 2964
Oy 641 LeuLeuLeuGlnLeuTyrlLeuGlnInTPGlyAspSerValLeuGlnArgGlyValLeu 660
Db 2965 CTTCCTGCTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3024
Oy 661 LeuHisSerGlnGlyAlaAlaSerSerValCysLysLeuAspGlyLeuLeuHisArg 680
Db 3025 CTGCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3084
Oy 681 PheLeuThrLeuLeuAlaAspThrSerAspSerArgAlaLeuGlnAspArgGlyAlaAsp 700
Db 3085 TTGATCAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3144
Oy 701 AlaSerMetAlaCysArgLysLeuAlaValAlaHisProLeuLeuLeuArgHisLeu 720
Db 3145 GGCACAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3204
Oy 721 PrometLeuAlaAlaLeuLeuHisGlyArgThrHisLeuAspPheGlnLysPheArgL 740
Db 3205 CCCATGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3264
Oy 741 GlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeuGlnLeuLeuGlnProHis 760
Db 3265 CAGAACCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3324
Oy 761 ValPheArgSerGlnHisGlnLysAlaLeuTrpAspCysLeuLeuSerPheLeuArgLeu 780
Db 3325 GTGTTCCGAGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3384
Oy 781 LeuLeuAsnTyrlArgLysSerSerArgHisLeuAlaAlaPheLeuLysPheValGln 800
Db 3385 CTGCTGAATTTACAGGAAGCTCCCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3444
Oy 801 PheLeuHisLysTyrlLeuThrTyrlAsnAlaProAlaAlaLeuSerPheLeuGlnLys 820
Db 3445 TTCATCCCAATATACATATCTACATATGCCAGCAGCCATCTCTCTGCGAGAGCAGC 3504
Oy 821 AlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeu 840
Db 3505 GCGGACCGCTCCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3564
Oy 841 LeuAlaGlyLeuSerLeuProSerArgAspAspArgThrAspArgLysLeuLeuGlnLys 860
Db 3565 CTTCGAGGGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3624
Oy 861 GlyGlnGlnGlnSerSerAlaGlySerLeuProLeuValSerValSerLeuPheLeuPro 880
Db 3625 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3684
Oy 881 LeuThrAlaAlaGlnMetAlaProTyrlMetLysArgLeuSerArgLysGlnLysValGln 900
Db 3685 CTGACCGCGCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3744
Oy 901 AspLeuLeuGlnValLeuSerAspLeuAspGlnMetSerArgArgArgProGlnLysLeu 920
Db 3745 GATCTGCTGAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3804
Oy 921 SerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGlnLysCysArgAsn 940
Db 3805 ACTTCTTCTGACCACTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3864
Oy 941 LeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSerLeuAlaAlaPhe 960

```

```

Db 3865 CTGCGCTTACGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3924
Oy 961 LeuProThrPheMetTyrlCysLeuGlnSerGlnAspPheGlnValGlnThrAlaLeu 980
Db 3925 CTGCCCCAGCTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3984
Oy 981 ArgAsnLeuProGlnTyrlAlaLeuLeuCysGlnGlnHisAlaAlaValLeuLeuHisArg 1000
Db 3985 CGGAACCTGCTGATGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4044
Oy 1001 AlaPheLeuValGlyMetTyrlGlnMetAspProSerAlaGlnLysSerGlnAlaLeu 1020
Db 4045 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4104
Oy 1021 ArgLysLeuHisMetGlnAlaValMet 1029
Db 4105 AGCATCTGATATGAGGCGCTGATG 4131

```

RESULT 3

```

LOCUS BC004286 4179 bp mRNA linear Pri 12-JUL-2001
DEFINITION Homo sapiens, clone IMAGE:3618123, mRNA, partial cds.
ACCESSION BC004286
VERSION BC004286.1 GI:13279124
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4179)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC/DCMP/DPF
 CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 BC Cancer Agency, Vancouver, BC, Canada
 info@bgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lissa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranda Tsai, Natasha van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 13 Row: 3 Column: 11
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

source

Location/Qualifiers

```

1..4179
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3618123"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH_MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOT7"
<1..2024
/codon_start=3
/product="Unknown (protein for IMAGE:3618123)"
/protein_id="AAH04286.1"

```

CDS

```

/db_xref="GI:13279125"
/translation="SRLGHILCRSRNPFLHIIORASSQSPWMLADVQSSSESL
DVLPGCLCEFLHDAADAGEDEDESKRQKRRQKROKOLGRLODLLE
PKADEQTECEVLDFLRILGSSQVARSVLAKMLSLSGSI.RDGEKPPMEEDYV
PDIVLOGYOMLRLDLPRLPLFDSRSTALALOOAHMEDIPTISNYLILYOHNPV
EBQAOHSDLDLVAARLVERSTTMSHLFSKLSASDAVLSLSIFSRVRRKOS
KBEVEYSWSESODVFLRWSSETATMHLVHAVALTLTLEPRADSEFOLDI
WPEEEKPLPTAFLVDTSSEALPLDWLKLMISEVLRVLDALLODLEPOULLFYOS
FGIPVSMKILQFLDQAVADPOTLEONIMDKNYMAHLEVQHEKASGGQTFHSL
TASLPBRDSTEAPKPSPEOPIGORIVGQLEVLAGEDELAGMFOIFPLSLDP
RMQSSRPVALALQOALGOELARVQSGSEVPGITRVGLATATLSSBEGALVMS
MRSRHLACPLRLQLOXORCVPRDITGFSSLEFKVLLOMQLDSDRVEGCPRLAOLR
MLASQASAGRRSDVGRGLRLKLEALAFRODLEVVSTVAFQPSPLMW"
BASE COUNT      712 a      1412 c      1337 g      718 t
ORIGIN
Alignment Scores:
Pred. No.:      4,53e-296      Length:      4179
Score:          4933.50      Matches:      976
Percent Similarity: 94.85%      Conservative: 0
Best Local Similarity: 94.85%      Mismatches: 2
Query Match:    94.13%      Indels:      52
DB:             9      Gaps:      1
US-09-929-769-7 (1-1029) x BC004286 (1-4179)
QY      1 MethisileuVal1HisAlaMetVal1IleuLeuThrLeuGlyProProArgAla 20
Db      912 ATGCACATTCCTCGTGGTCCATGCGCATGTCCTGCTGACGCGTGGGCGCCCGCCCGAGCC 971
QY      21 AspAspSerGluPheGlnAlaLeuLeuAspIleTrpPheProGluGluLysProLeuPro 40
Db      972 GACGACAGCAGATTCAGGCGCTGCTGACATCTGTTCCGGAGGAGAGCCACTGCC 1031
QY      41 ThrAlaPheLeuValAspThrSerGluGluAlaLeuLeuProAspTrpLeuLysLeu 60
Db      1032 ACCGCTTCTGCTGGACACATCGGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1091
QY      61 ArgMetIleArgSerGluVal1LeuArgLeuValAspAlaAlaLeuGlnAspLeuGluPro 80
Db      1092 CGCATGATCGGTTCTGAGTGGTCTCGGCTGGAGCGCGCGCCGCGAGGAGACTGAGCCG 1151
QY      81 GlnGlnLeuLeuLeuPheValGlnSerPheGlyIleProVal1SerSerMetSerLysLeu 100
Db      1152 CAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1211
QY      101 LeuGlnPheLeuAspGlnAlaVal1HisAspProGlnThrLeuGlnGlnAsn1IleMet 120
Db      1212 CTCACAGTTCCTGACACGAGAGTGGCCACGACCCCGACACTCTGGAGACAGACATCATG 1271
QY      121 AspLysAsnTrpMetAlaHisLeuVal1GlnVal1GlnHisGluArgGlyAlaSerGlyLys 140
Db      1272 GACAAAGATTAATACATGGCCACCTGCTGGAGTTCACGATGACGCGCGCTCCGAGAGGC 1331
QY      141 GlnThrPheHisSerLeuLeuThrAlaSerLeuProProArgArgAspSerThrGlnAla 160
Db      1332 CAGACTTTCACATCTTGGCTGACAGCCTCCCTGCGCCCGCCGACAGACAGACAGAGCA 1391
QY      161 ProLysProLysSerSerProGluGlnProIleGlyGlnGlyArgIleArgVal1GlyThr 180
Db      1392 CCCAAACCAAGACAGACCCAGAGACCCCATAGAGCCAGAGCCGCGATTCGGGTGGGAGCC 1451
QY      181 GlnLeuArgValLeuGlyProGluAspAspLeuAlaGlyMetPheLeuGln1IlePhePro 200
Db      1452 CAGCTCCGGGTGCTGGGCGCTGAGAGACCTGCTGCGATGTTCCCTCCAGATTTCGCCG 1511
QY      201 LeuSerProAspProArgTrpGlnSerSerSerProArgProVal1AlaLeuAlaLeuGln 220
Db      1512 CTGACGCTTGACCTCGGTGGACAGAGCTCCAGTCCCGCCCGGCTGCGCTGCGCTGAG 1571
QY      221 GlnAlaLeuGlnGlnLeuAlaArgVal1Val1GlnGlySerProGluVal1ProGlyTyr 240
Db      1572 CAGGCTTGGGCGCAGAGCTGGCGCTGCTCCAGGGCAGCCCGCGAGGTGGCGCATC 1631

```

```

QY      241 ThrValArgVal1LeuGlnAlaLeuAlaThrLeuLeuSerSerProHisGlyLysAlaLeu 260
Db      1632 ACGGTGGGTGCTCCGACAGGCGCTCGCACACCTGTCAGCTCCACAGAGGGGGTGGCTG 1691
QY      261 ValMetSerMetHisArgSerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGln 280
Db      1692 GTGATGTCACATGACACCTAGCCACTTCTGCTGCGCCGCTGCTGCGCCAGCTTCGACAG 1751
QY      281 TyrGlnArgCysVal1ProGlnAspThrGlyPheSerSerLeuPheLeuLysValLeuLeu 300
Db      1752 TACACAGCGCTGTGTCACAGACACCGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1811
QY      301 GlnMetLeuGlnTrpLeuAspSerProGlyVal1GlnGlyLysProLeuArgAlaGlnLeu 320
Db      1812 CAGATGCTGACAGTGGCTGGACAGACCTGGCGTGGAGGGCGGCGCCGCGGACAGCTC 1871
QY      321 ArgMetLeuAlaSerGlnAlaSerAlaGlyArgArgLeuSerAspVal1ArgGlyLysLeu 340
Db      1872 AGGATGCTTGGCACCCAGGCTTCAGCCGCGGCGAGCTCACTGATGTCGAGGGGGGCTC 1931
QY      341 LeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGlnVal1LysSerThrVal1 360
Db      1932 CTGCGCTTGGCGGAGGCGCTGCGCTTCCGTCAGAGCTCGAGGTGCTCAGCTCAGCT 1990
QY      361 ArgAlaVal1IleAlaThrLeuArgSerGlyLysGlnCysSerVal1GluProAspLeuIle 380
Db      1990 ----- 1990
QY      381 SerLysValLeuGlnGlyLeuIleGlnVal1ArgSerProHisLeuGlnGluLeuLeuThr 400
Db      1990 ----- 1990
QY      401 AlaPhePheSerAlaThrAlaAspAlaLysProPheProAlaCysLysProVal1Val 420
Db      1991 ----- CCGTTTCCAGCTGTTAGCCGCTTGTG 2017
QY      421 ValValSerSerLeuLeuLeuGlnGlnGluGluProLeuAlaGlyLysProGlyVal1Ala 440
Db      2018 GTGCTGAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2077
QY      441 AspGlyLysSerLeuGlnAlaVal1ArgLeuGlyProSerSerGlyLeuLeuVal1AspTrp 460
Db      2078 GACGCTGGACCCCTGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2137
QY      461 LeuGlnMetLeuAspProGluVal1Val1SerSerCysProAspLeuGlnLeuArgLeuLeu 480
Db      2138 CTGGAATGCTGGACCCCGAGGTGCTACGACGCTGCCCGACCTGACGCTGAGCTGCTC 2197
QY      481 PheSerArgArgLysGlyLysGlyGlnAlaGlnVal1ProSerPheArgProTyrTrpLeuLeu 500
Db      2198 TTTCCTCCGAGAGAGGCGCAAGGTCAGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2257
QY      501 ThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgVal1LeuLeu 520
Db      2258 ACCCTCTTCACGATTCAGTCCAGCTGACCTGCCACACTGACACAGTCCAGATTCCTGCT 2317
QY      521 GlyLysSerArgGlnArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCys 540
Db      2318 GCGAAGAGCGGAGAACAGAGTTCGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2377
QY      541 IleHisVal1ProArgIleTrpGlnGlyArgAspGlnArgThrProGlnLysArgArgGlu 560
Db      2378 ATCCATGCTTCTCGCATCTGACAGGAGGCGGAGACAGGACCGCCGAGAGCGCGGAGAG 2437
QY      561 GlnLeuVal1LeuArgVal1GlnGlyProGluLeuIleSerLeuVal1GlnLeuIleLeuAla 580
Db      2438 GAGCTGTGTGCTGGGCTCCAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2497
QY      581 GlnAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleGlnAlaArg 600
Db      2498 GAGCGGAGACCGCGAGACAGAGAGGAGGAGACAGCCGCTGACGCTTATTCAGGCGCGG 2557

```


D	b		1867	CTCCGTGCAGGGGCTCAGCGTGCACCAGACGAGGGAGCACAGACGCCAGACCAAGGCTTG :SANGA	1926
Oy			860	GluGluygluglUGUserSerAlaglYserIueProlEualSeValSerIuePhethr	879
D	b		1927	GAGGCGGAGGAGGAGAGCTCAGCGGCTTCCTGGCCCCCTGATCAGCGTCCCTGTTAC	1986
Oy			880	ProlEuPrAlaaIGUeMefAlaprobThymelytAsrgIueSerFarglyGIothVat	899
D	b		1987	CCTTGACCGCGGCGGAATGGCCCCCTTCATGAAGAAGCGTTTTCCCAGGAAAGCGGTG	2046
Oy			900	GUusPlaeLueGuVaIlleUsenSerSpIIeaSPJumetSerfARGARgProslule	919
D	b		2047	GAGATCTGCTGGAGACTCTGAGTAGACAATAAGACGAGATGTCCCAGGAGAACCCAGATC	2106
Oy			920	LueserPhepheserThrAsnLeugnInarIgneuwSetSereralagLUcLucysArsArg	939
D	b		2107	CTGACTCTCTCTCCACCACTGCAGCGGGCTGATGAGACTCGGCCGAGAGATTGTCGCG	2166
Oy			940	AstneuaIpapeserLeualaIeuunrsgernetiGaInsenseProSerIfelalala	959
D	b		2167	AACCTCGCCTTCAGCGTGCCTGGCCCTGCATGACGAAGAAGCCCAAGCATTCACACCGT	2226
Oy			960	PheLuePforthrPhemetyrcylsLeunglysersiGnasppheguIVaIVaISlnhra	979
D	b		2227	TTCCTGCCACAGTTATGTAAGCTGGGGACACCAAGACTTTGAGTGTGTCAGACGGCC	2286
Oy			980	LeuarGasnLeuProcLuTrArLaaleucycsyngIngIUHIsaIAlaVaIlleuLeuhis	999
D	b		2287	CTCCGGAACCTGCAGATAGCTCTCCTGNTCCAAAGACAGCGCGCTGCTGCTCAC	2346
Oy			1000	ArgAlaPheLueVaIGlymetetrLyginMeasApproserrAlacImIlesercGLua	1019
D	b		2347	CGGGCTTCCTGCTGGGATGTACGGCCAAGATGAGACCCCAAGCGCCAGATTCTCAGGCC	2406
Oy			1020	LeuarqileLueHisMetGLuaIVaIMet	1029
D	b		2407	CTGAGATCTGATATGAGAGCGCCCTGATG	2436
R	E	S	RESULT 5		
B	C	I	BC018777	2418 bp mRNA linear PRI 11-DEC-2001	
L	O	C	LOCUS	Homo sapiens, clone IMAGE:4869922, mRNA, partial cds.	
D	E	F	DEFINITION	BC018777	
V	E	R	VERSION	BC018777.1 GI:17511852	
K	E	N	KEYWORDS	.	
S	O	U	SOURCE	Homo sapiens.	
O	R	G	ORGANISM	Homo sapiens	
R	E	F	REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
A	U	T	AUTHORS	Strausberg,R.	
T	I	T	TITLE	Direct Submission	
J	O	R	JOURLNL	Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
R	E	M	REMARK		
C	O	M	COMMENT	NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: ATCC/DCTP/DTF CDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) BC Cancer Agency, Vancouver, BC, Canada info@bcsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letliccia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Matthewson, Candice McLesey, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeeji, Jacqueline Schlein, Duane Smilous, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy, George Yang, Scott Znyderduyn, Marco Marra.	

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 40 Row: k Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

SOURCE

```
source
1. 2418
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4869992"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIR_MGC_45"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
<1. 2097
CDS
```

CDS

```

/product="unknown (protein for IMAGE:4869992)"
/protein_id="TAH1877.1"
/db_xref="GI:17511853"
/translation="RSDVAGGGLRLAEALAFRODLEWVSTVRAVATYRLSGEOCSV
EPILISKVILGLIEVRSPHLELLTAFFSTADAPSPACPKVAVVSSLLDDEED
AGKPGADGGSIEAVRIGPSSGLIVMLEMDPEWVSSCPDIOILRFSRRKGGQAO
VSEFRPPLRLFTHQSSPPTLHOCITRLILCKSRQRPDPASDLIMACHTVPIWOG
RDORTPOKREEVILRVQGPESISVELTAAETRSQDGTACSLIOAQLPLLESC
CCDDSEVRVATLETSCIOQMGDSVIGRRCDLLOLYLORPELRAVVELLHISCEI
AASSVCKGLIHRTFTLLADSDSLANRKGDAACMKACVLAHPDLILLRLPML
AALLHGRTHLNFOEQRQNHLSCHLVHGLLELLQIPVFRSEHGALMDCLISTRILL
LNRKSSRHRLAATNKRVOEITKITYTNPAPASISLQKHADPLHDSFUNDVLVMS
LNLGILISPSDDRTDLGDEGEERSAGSLPVLSLFTPLTAEMAATYRLSRGCG
TYVDLLEVLSDIOMSRREPEISFSTNLIORLMSDECCRNATFALSWMQNSP
IAAEFLPTWYCKLGSQDFEVYQTLNRNLPETALLQGHANVLHRAFLVGMQMDPS
AOISEARLHMEAVM"

```

BASE COUNT
ORIGIN

Alignment Scores:

Pred. No.:	5,68e-211	Length:	2418
Score:	3554.00	Matches:	698
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	67.81%	Indels:	0
DB:	9	Gaps:	0

US-09-929-769-7 (1-1029) x BC018777 (1-2418)

QY	332	ArgLeuSerAspValAlaArgLgLYLeuLeuArgLeuAlaValAlaLeuAlaPheArgLn	351
Db	1	AGGCTCAATGATGTGCGAGGGGGGCTCTGGCGCTGGCGAGGGCTTGCTTCGTCAG	60
QY	352	AspLeuGluValValSerSerThrValArgAlaValAlaIleAlaThrLeuArgSerGlyGlu	377
Db	61	GACCTGGAGGTGGTCACGCTCCACCGTCCGTCCTCATCGCCACCCGTAGCTGTGGGAG	120
QY	372	GlnCysSerValGluProAspLeuIleSerLysValLeuGlnGlyLeuIleGluValArg	399
Db	121	CAGTGCACCGTGGAGCCGGACCTGTATGCAGAAAGTCCCTCAGAGGGGCTGATCGAGGTGAG	180
QY	392	SerProHisLeuGluGluLeuLeuThrAlaPhePheSerAlaThrAlaAspAlaAlaSer	411
Db	181	TCCCCCCACCTGGAGGAGCTGCTGACATCTTCTCTGCACCTGGGAGATGTCCTCC	240
QY	412	ProPheProAlaCysLysProValValValValSerSerLeuLeuGlnGluGluGlu	433
Db	241	CCGTTTCCAGGCTGTAAACCCGTTGTGTGTGTGAAGTCCCTGCGCTGCACAGAGAGAG	300
QY	432	ProLeuAlaGlyGlyLysProGluValAlaAspLysLysSerLeuGluGluValAlaArgLeuGly	451
Db	301	CCCTGCGCTGGGGGGAACCGGGGTGGGAGCGGTGGACCTGGAGGCGCTGGCTGGGG	360
QY	452	ProSerSerGlyLeuLeuValAspTrpLeuGluMetLeuAspProGluValValSerSer	477

Db 361 CCGCTGTCAGGCGCTCTAGTGGACTGGCTGGAATGCTGGACCCGAGCTGCTAGCACC 420
 QY 472 CysProAspLeuGlnLeuArgLeuLeuPheSerArgArgLysGlyGlnAlaGln 491
 Db 421 TGGCCCGACCTGCAAGCTCAAGGCTGCTCTCCCGAGAGAGGCAAGGTCAAGGCCAG 480
 QY 492 ValProSerPheArgProTyrLeuLeuThrLeuPheThrHisGlnSerSerTyrProThr 511
 Db 481 GTGCGCGCTGCTCCGCTCACTGCTGACCTGCTTCACGCAATCAGTCCAGCTGCGCCACA 540
 QY 512 LeuHisGlnCysIleArgValLeuLeuGlyLysSerArgLysGlnArgPheAspProSer 531
 Db 541 CTTCACACAGTGCATCCAGTCTGCTGGGCAAGCGGGAGACAGAGTTCGACCCCTCT 600
 QY 532 AlAserLeuAspPheLeuTyrPheAlaCysIleHisValProArgIleTyrGlnArgAsp 551
 Db 601 GCCTCTGTCGACTTCCCTCTGGGCTGATCCATGTTCTCTGCAATCTGGACGGGCGGAG 660
 QY 552 GlnArgThrProGlnLysArgArgGlnGlnLeuValLeuArgValGlnGlyProGlnLeu 571
 Db 661 CACGCGACCCCGAGAGCGGCGGAGAGCTGCTGCTGGGCTCAGGGCCGCGAGCTCTC 720
 QY 572 IleSerLeuPheAlaGlnLeuIleLeuAlaGlnAlaGlnThrArgSerGlnAspGlyAspThr 591
 Db 721 ATCAGCGCTGAGTGCATGCTGCTGGCCGAGCGAGACGCGAGCGGGAGACACA 780
 QY 592 AlaAlaCysSerLeuIleGlnAlaArgLeuProLeuLeuLeuSerCysCysGlyAsp 611
 Db 781 GCGCGCTGAGCGCTCATCCAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 612 AspGlnSerValArgLysValThrGlnHisLeuSerGlyCysIleGlnGlnTyrPheLysAsp 631
 Db 841 GATGAGAGTGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
 QY 632 SerValLeuGlnArgArgCysArgAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 651
 Db 901 AGCGTGTGGGCAAGCGGCTGCTGCGAGACCTTCTCTGCAAGCTCTCAACGCGCGCGAG 960
 QY 652 LeuArgValProValProGlnValLeuLeuHisSerGlnGlyAlaAlaSerSerSerVal 671
 Db 961 CTGGGCGTGGCGGCTGCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 QY 672 CysLysLeuAspGlyLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAspSer 691
 Db 1021 TGCAGAGTGCAGGACATCATCCACGCTTATCATCGCTTGTGGACACGAGGAGCTCC 1080
 QY 692 ArgAlaLeuGlnAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAlaValAla 711
 Db 1081 CGGGCGTGGAGAAACGAGGGGCGGATGCCAGCATGGCTGCGGAAAGCTGGCGGTGGCG 1140
 QY 712 HisProLeuLeuLeuLeuLeuArgHisLeuProMetIleAlaLeuLeuHisGlyArgThr 731
 Db 1141 CACCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 QY 732 HisLeuAsnPheGlnGlnLysPheArgGlnGlnAlaHisLeuSerCysPheLeuHisValLeu 751
 Db 1201 CACCTCACTTCCAGGAGGTTCGGGACAGAAACCATGAGCTCTTCTGCAAGTGTCTG 1260
 QY 752 GlnLeuLeuGlnLeuLeuGlnProHisValPheArgSerGlnHisGlnGlnAlaLeuTyr 771
 Db 1261 GGCGTGTGGAGCTGCTGCGACCCGACGCTGTCCGCGACGAGCAAGGGGCGCTGTGG 1320
 QY 772 AspCysLeuLeuSerPheIleArgLeuLeuLeuAsnTyrArgLysSerSerArgHisLeu 791
 Db 1321 GACTGCGCTTGTGCTTCAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 QY 792 AlaAlaPheIleAsnLysPheValGlnPheIleHisLysTyrIleThrTyrAsnAlaPro 811
 Db 1381 GCTGCGCTTCAACAAGTGTGTGCACTTCAATCAATCAATCAATCAATCAATCAATCAATCA 1440
 QY 812 AlaAlaIleSerPheLeuGlnLysHisAlaAspProLeuHisAspLeuSerPheAspAsn 831
 Db 1441 GCAGCGATCTCTCTGTCAGAAACGAGCGGCGGCTCCAGCAGCTGTCTCTGAGCAAC 1500

QY 832 SerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeuProSerArgAspAsp 851
 Db 1501 AGTACCTGGTGTATGCTGAATCCCTCTCTGCAAGGGGTCAAGCTGCGCCACAGGAGACAG 1560
 QY 852 ArgThrAspArgLysLeuAspGlnGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 871
 Db 1561 AGGACCGACCGAGCGCTGGAGAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 QY 872 LeuValSerValSerLeuPheThrProLeuThrAlaAlaGlnMetAlaProTyrMetLys 891
 Db 1621 CTGGTCAAGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
 QY 892 ArgLeuSerArgLysGlnThrValGlnAspLeuLeuGlnValLeuSerAspIleAspGln 911
 Db 1681 CGGCTTCCCGGGGCGCAAGAGGTGAGAGATCTGCTGAGGTCTGAGTGCATATAGACAG 1740
 QY 912 MetSerArgArgArgProGlnIleLeuSerPhePheSerThrAsnLeuGlnArgLeuMet 931
 Db 1741 ATGTCCCGGCGGAGACCCGAGATCTGAGCTTCTTCGACCAACCTGACGCGGCTGATG 1800
 QY 932 SerSerAlaGlnGlnCysCysArgAsnLeuAlaPheSerLeuAlaLeuArgSerMetGln 951
 Db 1801 AGCTGCGCGAGAGATGTTCGCGCAACCTGCTGCAAGCTGCGGCGCTGCTGCTGCTGCTGCTG 1860
 QY 952 AsnSerProSerIleAlaAlaAlaPheLeuProThrPheMetCysLeuGlySerGln 971
 Db 1861 AACAGCCCGACATTCGAGCGGCTTCTCTCCACAGCTTCAATGATCTGCGGCGAGCAG 1920
 QY 972 AspPheGlnValValGlnThrAlaLeuArgAsnLeuProGlnTyrAlaLeuCysGln 991
 Db 1921 GACTTGAAGTGTGGCAGAGGCGCTCGGAAACCTGCTGAGAGAGCTCTCTCTCTGCTGCTGCT 1980
 QY 992 GlnHisAlaAlaValLeuLeuHisArgAlaPheLeuValGlyMetCysGlnMetAsp 1011
 Db 1981 GAGCAGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
 QY 1012 ProSerAlaGlnIleSerGlnAlaLeuArgIleLeuHisMetGlnAlaValMet 1029
 Db 2041 CCCAGCGCGAGATCTCCGAGGCGCTGAGATCTGCTGATATGAGGCGGTGATG 2094
 RESULT 6
 LOCUS BC010333 2730 bp mRNA linear ROD 07-AUG-2002
 DEFINITION Mus musculus, clone IMAGE:360115, mRNA, partial cds.
 ACCESSION BC010333
 VERSION BC010333.1 GI:16307573
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2730)
 AUTHORS Strausberg, R.
 DIRECT SUBMISSION
 SUBMITTED (05-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT: MGC help desk
 EMAIL: cgabs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 CENTER CODE: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 CONTACT: amg@bcm.tmc.edu
 Guarantees, P.H., Garcia, A.M., Lu, X., Huiyk, S.M., Hale, S.M.,
 Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.


```

QY 692 ArgAlaLeuGlnAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAlaValAla 711
DB 1442 CGGCTCTTGGAGACAGAGTGGCTGATGCCAACATGCGCGGAGCGGCTGCGGCC 1501
QY 712 HisProLeuLeuLeuLeuArgHisLeuProMetIleAlaAlaLeuLeuHisGlyArgThr 731
DB 1502 CACCCCGCTCTGCTGCGAGGACCTGCCATGATGCGACACCTCTCGGACGCGGCGACT 1561
QY 732 HisLeuAsnPhenGlnGluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeu 751
DB 1562 CATCTAAACTTCCAGGAGTTCGCGGACAGAACACCTGCGCTTCTGTCATGTCCTG 1621
QY 752 GlyLeuLeuGluLeuGlnPheArgSerGlnHisGlnGlyAlaValLeuTrp 771
DB 1622 GGCATCTGGAGCTGCGAGCGCGGCTGCTCCAGAGCAACACCGGGGGGCTTGG 1681
QY 772 AspCysLeuLeuSerPheIleArgLeuLeuLeuAsnArgLysSerSerArgHisLeu 791
DB 1682 GACTGCGCTGCATCTTATCCGCGCTGCTCCTGACACTACCGGAAGTCTCCGCGCACTG 1741
QY 792 AlaAlaPheLeuAsnLysPheValGlnPheIleHisLysTyrIleThrTyrAsnAlaPro 811
DB 1742 GCTCCCTTCTTAGCAAGTTCGTTCAAGTTCATCCACAACTACGTGGGCTCCAGCGCACCA 1801
QY 812 AlaAlaIleSerPheLeuGlnLysHisAlaAspProLeuHisAspLeuSerPheAspAsn 831
DB 1802 GCAGCGGTGGCTTCTCGCGAAGACGACGAGCGCCCTCATGACCTGCTCTTGGACAC 1861
QY 832 SerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeuProSerArgAspAsp 851
DB 1862 AGTACCTGCTGATGTTGAGAGTCTTACTGCGAGGCTCAGCGCCGCGGATGGG 1921
QY 852 ArgThrAspArgGlyLeuAspGluGluGluGluGluGluSerAlaGlySerLeuPro 871
DB 1922 AGACTGACCAAGCGCTGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1981
QY 872 LeuValSerAlaSerLeuPheThrProLeuThrAlaAlaGluMetAlaProTyrMetLys 891
DB 1982 CTGCTGAGTGTCTCCCTCTCCACCCCATGACTGATGACCCAGCGGCTCTCATCATAG 2041
QY 892 ArgLeuSerArgGlyGlnThrValGluAspLeuLeuGluValLeuSerAspIleAspGlu 911
DB 2042 AGCTGCTCCCGGGGCGGCGGCTGTTGAGAGTGTCTGACAGCCTGAGTGCATGACAGAG 2101
QY 912 MetSerArgArgArgProGluIleLeuSerPhePheSerThrAsnLeuGlnArgLeuMet 931
DB 2102 ATGTGCGGGGAGACCTGAGTCTGGGCTTCTTCTGACCAACTTGGACGGCTGATG 2161
QY 932 SerSerAlaGluGluCysCysArgAsnLeuAlaPheSerLeuAlaLeuArgSerMetGln 951
DB 2162 AGTCAAGCGGAGAGTCTCCGCAACCTGCTTACAGCTTACAGCTGCGCTCATTCAG 2221
QY 952 AsnSerProSerIleAlaAlaAlaPheLeuProThrPheMetTyrCysLeuGlySerGln 971
DB 2222 AACAAACCCGACATCGCTGCTGCTGCTCCACACTTCACTGACTCTGCGGACAGAGA 2281
QY 972 AspPheGluValValGlnThrAlaLeuArgAsnLeuProGluTyrAlaLeuLeuCysGln 991
DB 2282 GACTTGGAGTGTATACAGACGCTTCAAGACCTGCGCAAGTACACCTGCTGCGCA 2341
QY 992 GluHisAlaAlaValLeuLeuHisArgAlaPheLeuValGlyMetTyrGlyGlnMetAsp 1011
DB 2342 GACACGCGGCTGCTGCTGCGACAGAGCTTCTTGTGGGCGGTACGGCCAGATTGAC 2401
QY 1012 ProSerAlaGlnIleSerGluAlaLeuArgIleLeuHisMetGluAlaValMet 1029
DB 2402 ACCAGCGCCAGATATACAGAGCCCTGACATCTCAACATGAGGCTGTGATG 2455

```

RESULT 7
 HSM800197
 LOCUS HSM800197 2241 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA; cDNA DKFZp586J0619 (from clone DKFZp586J0619);

```

ACCESSION partial cds.
VERSION AL050110
KEYWORDS AL050110.1 GI:4884139
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS 1 (bases 1 to 2241)
TITLE Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
JOURNAL Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOMA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp586J0619) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
source
1. 2241
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp586J0619"
/tissue_type="uterus"
/clone_lib="586 (synonym: hute1). Vector pSport1; host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
1. 1924
/gene="DKFZp586J0619"
<1. 1924
/gene="DKFZp586J0619"
/feature="unknown"
/note="unknown"
/codon_start=2
/product="hypothetical protein"
/protein_id="CA843378.1"
/db_xref="GI:4884140"
/db_xref="SPTREMBL:Q9Y3W8"
/translacion="TSSRSRAELTATFFSATADASPFPACKRVVYVSLLOEEBPL
AGKPGADGSLAVRLGPSGLVLMLEMDPEVYSSCPDQLRLFSRKGKGOAO
VSEFRYLTLTFHOSMPTLHOCIVLLKASEORFDPDSASIDFLMACIHVPRIWG
ROORPORKREELVAVRGPELISLVEILAEKTRSGDGTACSLIARPLILSC
CGDDSVKVTYHESGCIQOMGDSVLRGRRLQIQLQRELPVPPVYVLSSEG
AASSVCKUDGLHRTTLADTSRALLENKADSMCKLVAHPILLHLHPWI
AALLHGRHLNDFEQONHLSCLFLVLELLELOHVRSEHOGALWDCLSFRL
LNYRKSSRLAAFINKEVOFIRKYITVYNAPAISFQKADPLHDSFNSDLYMKS
LLAGLSLPSRDRDTGDLDEEGEESASGLPVSFLTPITAAEMAPYMKRLSRGO
TYEDLLEVLSDIDEMSRRRPELISFSTNLIORIMSSAECCRLARSLALRSQNSPS
IAAARLPTPYWCKIGSDPEFVYVOTALNLPYALILCOEHAAVLLHRAFLVGMQMDPS
AQISEALRLIHMAVAV"
BASE COUNT 383 a 754 c 712 g 392 t
ORIGIN
polya_site
2224
2224
Alignment Scores:
Pred. No.: 8.04e-192 Length: 2241
Score: 3244.00 Matches: 633
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.90% Indels: 0
DB: 9 Gaps: 0
US-09-929-769-7 (1-1029) x HSM800197 (1-2241)
QY 397 GluLeuLeuThrAlaPhePheSerAlaThrAlaAspAlaSerProPheProAlaCys 416
DB 23 GAGCTGCTGACGATCTTCTCTGCGACATGCGGATGCTGCTCCGCTTCCAGCCTGT 82
QY 417 LysProValValValSerSerLeuLeuGlnGluGluGluProLeuAlaGlyGly 436

```


JOURNAL Submitted (23-NOV-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 148996)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (22-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 148996)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 6 (bases 1 to 148996)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 7 (bases 1 to 148996)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Jun 1, 2002 this sequence version replaced g1:19774560.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sepiens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH1246C19

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7/>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-369D24, 2000 bp overlap; the clone sequenced to the right is RP11-16P10, 2000 bp overlap. Actual start of this clone is at base position 174583 of RP11-369D24; actual end is at base position 63205 of RP11-16P10.

Sequence derived from one plasmid subclone, base position 32899 to 33087.
 FEATURES
 source
 Unresolved tandem repeat from base position 99549 to 100076.
 1..148996
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"
 /clone="RP11-1246C19"
 /clone_1lb="RPCI-11"
 585..893
 /rpt_family="Alu"
 repeat_region
 2518..2618
 /rpt_family="MIR"
 repeat_region
 2708..2769
 /rpt_family="MIR"
 repeat_region
 2873..3034
 /rpt_family="MIR"
 repeat_region
 3347..3438
 /rpt_family="MIR"
 repeat_region
 3439..3574
 /rpt_family="L1"
 repeat_region
 3575..3857
 /rpt_family="Alu"
 repeat_region
 3858..4034
 /rpt_family="Alu"
 repeat_region
 4035..4628
 /rpt_family="L1"
 repeat_region
 4629..4714
 /rpt_family="Alu"
 repeat_region
 4796..5037
 /rpt_family="Alu"
 repeat_region
 5039..5296
 /rpt_family="Alu"
 repeat_region
 6009..6933
 /rpt_family="L1"
 repeat_region
 6938..7061
 /rpt_family="Alu"
 repeat_region
 7145..7243
 /rpt_family="L1"
 repeat_region
 7256..7559
 /rpt_family="MER1_type"
 repeat_region
 9418..9677
 /rpt_family="Alu"
 repeat_region
 9758..9876
 /rpt_family="L2"
 repeat_region
 9886..9926
 /rpt_family="MER1_type"
 repeat_region
 9927..10207
 /rpt_family="TTCC)n"
 repeat_region
 10208..10220
 /rpt_family="Alu"
 repeat_region
 10227..10366
 /rpt_family="MER1_type"
 repeat_region
 11484..11545
 /rpt_family="L2"
 repeat_region
 12151..12259
 /rpt_family="MIR"
 repeat_region
 12792..13111
 /rpt_family="MIR"
 repeat_region
 14332..14451
 /rpt_family="Alu"
 repeat_region
 14794..14822
 /rpt_family="MIR"
 repeat_region
 14828..15133
 /rpt_family="AT-rich"
 repeat_region
 15194..15288
 /rpt_family="Alu"
 repeat_region
 16148..16439
 /rpt_family="MIR"
 repeat_region
 /rpt_family="Alu"

```
repeat_region 16468..16637
                /rpt_family="Alu"
repeat_region 16647..16944
                /rpt_family="Alu"
repeat_region 16972..17244
                /rpt_family="Alu"
repeat_region 17254..17442
                /rpt_family="L1"
repeat_region 19130..19311
                /rpt_family="L2"
repeat_region 19432..19683
                /rpt_family="MIR"
repeat_region 19833..19938
                /rpt_family="L2"
repeat_region 20105..20243
                /rpt_family="MIR"
repeat_region 20951..21003
                /rpt_family="L1"
repeat_region 21298..21415
                /rpt_family="L1"
repeat_region 21589..21644
                /rpt_family="MIR"
repeat_region 21944..22036
                /rpt_family="MIR"
repeat_region 22180..22292
                /rpt_family="MIR"
misc_feature 22268..23087
                /note="CGC island (%GC=72.4, O/e=0.80, #CpGs=75)"
repeat_region 22660..22845
                /rpt_family=" (CACCC)n"
repeat_region 23254..23354
                /rpt_family="MIR"
```

Alignment Scores:	
Pred. NO.:	1 41e-186
Score:	3199.50
Percent Similarity:	24.85%
Best Local Similarity:	24.83%
Query Match:	61.05%
DB:	9
	Gaps: 22
	Matches: 148996
	Conservative: 1010
	Mismatches: 1
	Indels: 6
	Gaps: 3054

US-09-929-769-7 (1-1029) x AC102953 (1-148996)

QY	13	LeuhtlrleuqltPro-----	-Proatqalaasp	21
Dp	136178	TTGGGCGTGGCCCCCGGTGGGTCTGAAGCCTGCATCCGCTTCCTTCCTTGAAAGCCGAC	 	136119
QY	22	AspergerGlutPheglinalaleuLeuAspIlaetRPhetroclustluySProleuProthr	 	41
Dp	136118	GACAGCCAGATTCCAGGGCGTCTGGACATCTGGTTTCCGAGGAGAAGACCACAGCCGCCAC	 	136059
QY	42	AlaphelauValasprThrserGlutlualaleuLeuProksprThpleuylsLeuarq	 	61
Dp	136058	GCCTTCCCTGGGTGGACACATCGAGAGAGGCCCTGCTGCTTCTCATCGACTGGAGCTGGCG	 	135999
QY	62	MettlearSerSerGuValleuArgLeuValAspIla-----	 	73
Dp	135998	ATGATCCGTTCTTAGGTGCTCTCCGCTTG7GAGACGCCGCGTACAGGCTCAAGCTGGGCGTTGGG	 	135939
QY	73	-----	 	73
Dp	135938	CAGGGGCTGCTGACTCAAGGGGTCCCAGGGCCCCCTTGCTGCTCCTTGtGGGCTTGCGGAC	 	135879
QY	73	-----	 	73
Dp	135878	CTCCGGGCTGTGGCAGAGAGAGAAGACTGGCCGGGTAGCCTCAAGGGGAATCCTGCAGACCATTA	 	135819
QY	73	-----	 	73
Dp	135818	TGTCTTGCCACACAGAGCTGTGCTGAGACCTTGAGAGAGAACCMAGGGCACGCTGTGCTG	 	135759
QY	73	-----	 	73

Db 135758	GCCTCGCCGAGAGGGTCTTACTGTCAAGGCTCAGGGGTCTAAGTGTAGAAACAGGTCCTC	13569
QY	73	73
Db 135698	GATGAAAAAAGAGAAAGTGTCTGGGGTTGGAGAGGAGAGAGAAGTGGTGGGCTTC	135639
QY	73	73
Db 135638	TGGGTGGCTCCCTGAAAGATTGGGGGTTGGAGCTGCAGCCGGGGAGCCCGACAAGACT	135579
QY	73	73
Db 135578	GAGCAAGGGGCTGGGCCCAAGTACGTACGTAGAGGGCTGTCTGCTCACAGTGAAGAGA	135519
QY	73	73
Db 135518	CACGTGGGGGACAGCTTGTGAGTGAATGGGCTGGCTCTGAAACCATGACTGCCAAGG	135459
QY	73	73
Db 135458	GTTAGGCTGTTGTGTGTGCCCGGAAGCCACCATGACTTCGGGGCAAGAGCCCTTGAG	135399
QY	73	73
Db 135398	CACCTAGAGAAATATGTACACAGCAGCTGCTGCAGCTTGGAAGTACTGGCCGTGC	135339
QY	73	73
Db 135338	CACCATAGGCCCTTGAGAGCGCCACCCACCCCTGGGGATGGGGCCACACAGCTGCAT	135279
QY	73	73
Db 135278	CCAGGGTGCAAGTGTGTCCGACACGCTGCCCTGTGGGGCCCAAGTCTCGGGTGG	135219
QY	73	73
Db 135218	GTACACGAGCCACGTGGGCCAGGGTCTCAGGTGGGGCTGCCGAGTTCGACTGCCACG	135159
QY	73	73
Db 135158	GTCTGGGTGGGGCTGCGGAGTGTCACTGCCCCAGGCTTGGGTGGGTACACGAGCC	135099
QY	73	73
Db 135098	ACGTGCCCCAGGGTCTCAGTGGGGCTGCCGAGTGTGCACGCCACAGTCTTGGGTGG	135039
QY	73	73
Db 135038	GTACCCGAGCCACGTGGGCCAGGGTCTCAGTGGGTGTCACTCACACAGCAGCC	134979
QY	73	73
Db 134978	CCAGGTCTCGGGTGGGGTCCGAGTGGCCATTGCCCGGGTTTGGGTGGGCTTAAGGAGC	134919
QY	74	89
Db 134918	GGGTGCACCTCAGCCCTGCAGGACTGCAGCCGAGAGCTGCTGTTCTGTCGACATC	134859
QY	89	109
Db 134858	GTTTGGCATCCCCGTGTCCAGCATGAGCAAACTCTCCAGTTCTGAGCAGGCAGTGGC	134799
QY	109	122
Db 134798	CAACGACCCCCAGACTCTGGAGCACAACATCATGAGACAAAGATAAGATGCCAGGGACCTG	134739
QY	122	122
Db 134738	CGAGATGGGTTCACACACGCTGACCGAGCCCGAGCCCGGACGCTTTGGGCTGGGT	134679
QY	122	122
Db 134678	GCGAGCGGGTGGGTGATCCGAGGTCGGGTGGGTGTGGGTGGCACTCGAGGGCAGGCTTAG	134619

QY 122 ----- 122
 Db 134618 ACCTGTGGTGTGAGGCTGACAGTCAGAGGAGCGGAGCCCTGCGCCCAA 134559
 QY 122 ----- 122
 Db 134558 CCGTGGGGCCAGGACCCCTAGGGCCGTGTTTGTGAGGAGCTGTGGGGCCT 134499
 QY 122 ----- 122
 Db 134498 CACTGCTGACTGACTCACCACCTTCTGAAGCTTTGGGAAGTGTGTGAGGTGGG 134439
 QY 122 ----- 122
 Db 134438 GCATCTCAACTCTGGTTTCTGTGAGAGGAGAGACCTGCTCCCTGGCCCTGGCCGG 134379
 QY 123 ----- 133
 Db 134378 CCCAGGTGTGACCTCTCTACGGTTTGCAGATTACATGGCCACCTGTGGAGGTCCAGCA 134319
 QY 133 sgluArglyAlaSerGlyGlyInthrPheHisSerLeuThrAlaSerLeuPropr 153
 Db 134318 TGAGCCGGGGCCCTCCGGAGGCCAGACTTTCACCTCTCTCAGACCTCCCTCCCGCC 134259
 QY 153 oArgArg ----- 155
 Db 134258 CCGCCAGAGTACTGGGGAGGGGTGGGGCCAGCACAGAGACCAGCTGGGGAGGGGTGGG 134199
 QY 155 ----- 155
 Db 134198 GCCACACAGAGACCCAGCTGGGGAGGGGGAGCGGGCCCTGCACAGCCACACAGGCGAGT 134139
 QY 155 ----- 155
 Db 134138 GCACACTTGTGATGCACTAGAGAGGGGCCCCACAGAAACTCTCAGCTGTCTTTGGG 134079
 QY 155 ----- 155
 Db 134078 CAGCAATTCCTCTGTGGCAGCATGACAGTTCCTTCAGAGTTCAGATCCCTGTGGAC 134019
 QY 155 ----- 155
 Db 134018 CTGAGGAGACCTGGGGTCTTGGGAGAACACAGCCAGCTGTGAGCTTCTCTCTGG 133959
 QY 155 ----- 155
 Db 133958 GTGACAGTCCCGGGGAGAGACTGCGCATGGCTGTACAGTTCACCGCTGGAAGCGAGT 133899
 QY 156 ----- 162
 Db 133898 CAGAAAGTGTCCAGCAGTACAGTCTCTCTCTGTGACAGCAGCAGAGCCACCCAA 133839
 QY 162 sProLysSerProGluInProIleGlyInGlyArgIleArgValGlyThrGlnLe 182
 Db 133838 ACCAAAGAGAGAGCCAGAGAGAGCCCATAGGCCAGGGCCGATTCGGGGTGGGAGCCAGCT 133779
 QY 182 uArgValIleuGlyProGluAspAspLeuAlaGlyMetPheLeu ----- 196
 Db 133778 CCGGTGTGTGGCCCTGAGAGACGACCTGGCTGTGATGTTCTT - CCAAGTAGCGTTGCTGC 133720
 QY 196 ----- 196
 Db 133719 TGCATGTGGGTCTCACCTGACAGCCCTCTCTGTGCTGAGAGCTTCGCTGCGCCTCGC 133660
 QY 196 ----- 196
 Db 133659 AGCCACAGCGGCTTACTTGTGCTGACCGCTTACCCAGGACGCTGACAGCAGCAGCG 133600
 QY 196 ----- 196
 Db 133599 GCCTTTACAGCGCTGACCGCTGACCCCTGCGACGCTGCTTTTACAGTAACACCCCATGCT 133540

QY 196 ----- 196
 Db 133539 CTCAGATGAGAGCTAGGCCAACATTGGGAGAGGGGTCCGCTCATCTCTCTGCCATC 133480
 QY 196 ----- 196
 Db 133479 CTTGCACACCCCTCTCGCGGGGGCCACCCGGTTTGATGTGTACAGAACTGAGGTTCCAGG 133420
 QY 196 ----- 196
 Db 133419 TCTCGGCTGAGTCTGTGGCTGTGGAGCCGGGGATCTCTGCAGGCTCAGCACCTCTGATG 133360
 QY 196 ----- 196
 Db 133359 GAGAGCTGTGTGCTCCCTGGGAGAGCTGTGACACCTTTGTTCTGTGGGCACTCTCTGG 133300
 QY 196 ----- 196
 Db 133299 GTGTTGGGGTCCCTCACTGCTTCACATCACAGGTCTCAAGGGCCCTGGGCCCTCAGGCC 133240
 QY 196 ----- 196
 Db 133239 ACAAGGTGGGGCTACGAATGAACAGAGGACAGATGGAGCGGGTTCTGGCAGGAGCTG 133180
 QY 196 ----- 196
 Db 133179 CACCGGGCGCTTAGCCCTTCGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTGACA 133120
 QY 197 ----- 211
 Db 133119 CTGGCCCTCCCGCTTCACAGATTTCGGCTCACGCCCGGACCCCTCGGTGGCAGAGCTCCAG 133060
 QY 211 rProArgProValAlaLeuAlaLeuInGlnAlaLeuGlyInGluLeuAlaArgVala 231
 Db 133059 TCCCGCGCCCGTGGCCCTGCGCCCTGCAGAGGAGCTGTGGGACAGAGACTGGCCGCGCTCG 133000
 QY 231 IGIInGlySerProGluValProGlyIleThrValArgValLeuGlnAlaLeuAlaThrLe 251
 Db 132999 CCAGGGCAGCCCGGAGGTGCGGGGACATCAGGTGCGGTCTCTCAGAGCCCTCCACCCCT 132940
 QY 251 uLeuSerSerProHisGlyGlyAlaLeuValMetSerMetHisArgSerHisPheLeuAl 271
 Db 132939 GCTCAGCTCCCGCACAGCGCGCGTCCCTGTGATGTCCATTCACACCTGACACTCTCTGGC 132880
 QY 271 acysProLeuLeuArgGlnLeuCysGlnTyr ----- 281
 Db 132879 CTGCCCGCTCTGCGCAGCTGTGACAGTACTTGGGCGTGGGCTCTCTCTGCGG 132820
 QY 281 ----- 281
 Db 132819 GAGAGAGGCGATGAGGGCTCTGCTCCGCGAGAGCGCCGGGTCTCTGTGACTGCTC 132760
 QY 281 ----- 281
 Db 132759 CTCTGTGGGTCTTGGGTCTCAGTTTCCAGTAGTAAGAAAGCACTGGCCCAAGGAGTGGG 132700
 QY 281 ----- 281
 Db 132699 GCAGTCCCTTTTTCAGTAGAATTCCCGGAGAGGGAGGTGGGGAGTGGCGCGT 132640
 QY 281 ----- 281
 Db 132639 CCGCTGTGTTCACACCAACAGGCTGCGGCTCTGCTGCTTCTCTTCCCATCTGCTGTAC 132580
 QY 281 ----- 281
 Db 132579 CTAGGACTGTGTTTACTTGTGTCTGTGTACTTTTCTTCTTGAACCGCGTGGAGGCC 132520
 QY 281 ----- 281
 Db 132519 CTCACAGCTGGGGGGTTTGAAGAGGTCTGTGAGTGTCTGTGGGCCGTGACAGGGGA 132460
 QY 281 ----- 281

Db 132459 GGCACCGTTTCAGACAGAGAGCGGGATGAGTTGGCCACGGCTCCAAAGTCCCGTGGCG 132400
 QY 281 ----- 281
 Db 132399 CAGGTCCTCAGACGCTCAGGGGCCCTGCTGGGCTCGGTCGCCGCTCCCGTGTCTC 132340
 QY 282 ----- 282
 Db 132339 CTCTGCCAGGGTAGCCACCTGTCCTTTGCTCCCGCAGCGCTGTGTGCCACAGACAGCG 132280
 QY 290 lYpHeserSerleuPheLeuValLeuLeuGlnMetLeuGlnTrpLeuAspSerProG 310
 Db 132279 GCTTCCTCGCTCTCTCTGAAAGTGTCTCTGAGATGCTGTGAGTGGTGGACAGCCCTG 132220
 QY 310 lYValGluGluGlyProLeuValGlnLeuArgMetLeuAlaSerGlnAlaSerAlaG 330
 Db 132219 GCGTGAAGGGGGGGCCCTGCGGGGCACAGCTCAGATGCTTGCCACAGCGCTCAGCCG 132160
 QY 330 lYArgArgLeuSerAsp ----- 335
 Db 132159 GCGGCGAGGCTCAGTGTGTGTGAGAGACGGCGCGGCTGGCAGAGCGCTGGGGGCACTGG 132100
 QY 335 ----- 335
 Db 132099 ACAGGGGCTCAGTCAGAGTGGCCAAAGCTGCTGGCACCCCGAGGCTGGGAGTGGCTTG 132040
 QY 335 ----- 335
 Db 132039 GGGGGTGGGTACCCCGCTGGGGCCAGGGAGCCCTTCTACCTGGGGGCTCCCTGATGG 131980
 QY 336 ----- ValA 337
 Db 131979 GATACACCGATGGGGTGAGGGCTGGAGCCACAGGGGCACTCAGACACACTCCACAGTGC 131920
 QY 337 rGlyGlyLeuLeuArgLeuAlaGlnAlaLeuAlaPheArgGlnAspLeuGlnValValAs 357
 Db 131919 GAGGGGGGCTCTCGCCCTGGGGCGAGGCCCTGCTCCGTGAGAGCTGGAGGTGTGCA 131860
 QY 357 eSerThrAlaArgAlaValAlaLeaThrLeuArgSerGlyGlnGlySerValGluP 377
 Db 131859 GCTCACCCCTCGTGGCTGCATCGCCACCTCAGTGTGGGGAGAGTGCAGGTGGAGC 131800
 QY 377 rAspLeuLeuSerLys ----- 382
 Db 131799 CGGACCTGATCAGCAAAAGGTACCTTCCTGGCACATCCTCTTGGGGCAGGGGCCCTGG 131740
 QY 382 ----- 382
 Db 131739 GAAGACCAAAACCTGAGGCTCTCTTAGCTGCCAGTGGGGAGAGGAGTGTCTCTCA 131680
 QY 382 ----- 382
 Db 131679 GCAGTGTCCCGGAAAGTCCACAGAGGTCCCTCCAGATGGCTCAGGGCGGCGCTTC 131620
 QY 382 ----- 382
 Db 131619 TTGAGAAATGATGTGAGTCTCTCTTGGGGGGGCCCTGATCACTGCTGCACCCAGCAAA 131560
 QY 382 ----- 382
 Db 131559 GTAGAGCCCGGGCGCGGCGGCACTTCAGACCCCGGCGCAGCTCTCTCCGCGGCCCGG 131500
 QY 382 ----- 382
 Db 131499 CTGCACTCTCTCTGTGCTCGGCTGAGACTGTCTTGTGTCCACACATGCTTCACCC 131440
 QY 382 ----- 382
 Db 131439 CACACCCCACTCGCTGAGCCACACCCCACTAGAGGGGCCCAACCCGAGCTGGAGATGAC 131380
 QY 383 ----- ValLeuGlnGlyLeuIleGluValArgSerProH 394
 |||||||

Db 131379 CCACATGTCCTTTGTCCCTGGCGAGTCCCTCCAGGGGGTATGAGTGAAGTCCCCC 131320
 QY 394 lSLeuGluGluLeuLeuThrAlaPhePheSerAlaThrAlaAspAlaIleSerProPhe 414
 Db 131319 ACCTGAGAGAGCTGCTGATGCAATTCCTCTGCTCCAGCTCGAGTGTGCTCCCTTC 131260
 QY 414 rAlaCysLysProValValValIleValSerSerLeuLeuGlnGluGluProLeuA 434
 Db 131259 CAGCTGTAAAGCCCGCTGTGTGTGTGTAGCTCCCTGCTGTGCAGAGAGAGAGCCCTGG 131200
 QY 434 lAeGlyLysProGluAlaAspGlyLys ----- 443
 Db 131199 CTGGGGGAAAGCCGGGTGGGAGAGGTGGCAGAGTGAAGTGGGGGTGTGCACAGCGGAGG 131140
 QY 444 ----- SerLeuGluAla 447
 Db 131139 TGGGACGGCGGGTGGACACAGCCCTTCCTGAGCCACTTCTCCAGACCTGGAGGGC 131080
 QY 448 ValArgLeuGlyProSerSerGlyLeuLeuValAspTrpLeuGlnMetLeuAspProGlu 467
 Db 131079 GTGGGCTGGGGGCTCGTCAGGCTCTCTAGTGTGACTGGCTGGAATGCTGGAGCCCGAG 131020
 QY 468 ValIleSerSerCysProAspLeuGlnLeuArgLeuPheSerArgLys ----- 485
 Db 131019 GTGGTCAGAGCTGCCCCGACCTGACGCTCAGGCTCTCTCTCCCGAGGAGGTGGGT 130960
 QY 485 ----- 485
 Db 130959 GGCACCCAGCTTCAGACAGGGGAGCCACCCAGGGTAAGGCTTCTCTGACTGAGTGG 130900
 QY 485 ----- 485
 Db 130899 CAGCAGCTTGTGGCTGGGGAGGCTTGCAAGGGGCCAGACCCCGCTGTCAAGGC 130840
 QY 485 ----- 485
 Db 130839 ATGTGTCCCTGTAGCAACCCAGAGAAAGGTAGTCCCCCAATGGCTTAGCAGAAACCA 130780
 QY 485 ----- 485
 Db 130779 GGCAGGGTCACAGTCACTGTGGACAGCAGTGTGACCAAGCCAGATGTGGGATGGGT 130720
 QY 485 ----- 485
 Db 130719 GGGTGAATGGGGTGGGGCTGGCCCTGCAAAACCTGTAGATCATGTAGGGCCCTCATGC 130660
 QY 485 ----- 485
 Db 130659 TCCCAAGGGAATGGGGCTGTGGCCAGAGAAAGGAGGGAGGTGCTGGCTCCACAT 130600
 QY 485 ----- 485
 Db 130599 CCTCAGAGGAGGAGCGGCTCACCAGAGAGGCTGATGGAGCCCGGGCACTCCCTGT 130540
 QY 485 ----- 485
 Db 130539 GTGGCATGGCTCCCGTGGGAAGGGGGGGAGATGCAAGCCAGCAGCAGAGAGGGCCCT 130480
 QY 485 ----- 485
 Db 130479 CCAGCAGATGAGGGGGCGGTGTATACGGCAGCAGCAGCAGCAGCAGCAGCAGCAG 130420
 QY 485 ----- 485
 Db 130419 CTCATTGGGGAAGCTGTGAGTGTGAGCAGAGAGCTGCCATTGGCCAGGTGAGGCA 130360
 QY 486 ----- GlysGlyGlnAlaGlnValProSerPheArgPr 497
 Db 130359 TCCGGGCTTCTCTCCCTCCCGCAGGGCAAGGTCAAGGCCAGGTCCCTGCTCGTCC 130300
 QY 497 oTyrlLeuLeuThrLeuPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleAr 517
 Db 130299 CTACCTCTGACCCCTCTCAGCATGATGTCAGCTGGCCACACATGTCAGCAGTGCATCCG 130240

Db 128081 TAACAGGGGGTCTGCGGCTCTCCAGCCTGGGCAACAAGACAATACTCCAAAAACAAC 128022
 QY 717 ----- 717
 Db 128021 AAAAAAGCATGTTCCATGGGCTGAGCTTGACCGGGGGAAGTACGAGCTGAGCTGGGC 127962
 QY 717 ----- 717
 Db 127961 CTCAGGTGGAGAGAGCTGGGTGGGCGCTCAGAGAGGGGATCTCATCCCACTCTCTAG 127902
 QY 717 ----- 717
 Db 127901 GGCAGAGAGGTGTGGGGGGCAGCAGCGGCTCTGCTCCGGGGCAGAGAGTGGGCGCT 127842
 QY 717 ----- 717
 Db 127841 TCTGGGGAGCTGGCCGAGGTGAGCCGCGGGGGGCTGACTGGCCCTCACCGCGCC 127782
 QY 718 ---ArgHisLeuProMetLeuAlaLeuLeuHisGlyArgThrHisLeuAspPheGln 736
 Db 127781 CGCAGGCACTGCCCCATGATGCGGCGCTCTCGACGGGCGCCAGCTCACTTCAG 127722
 QY 737 GluPheArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGlyLeuLeu3Leu 756
 Db 127721 GAGTTCGCGCAGCAAAACACCTGAGCTGCTTCCTGACAGCTGTGGCGCTGCTGAGCTG 127662
 QY 757 LeuGlnProHisValPheArgSerGlnHisGlnGlyAlaLeuThrAspCysLeuLeuSer 776
 Db 127661 CTGACGCGCAGCTGTCTCCGAGCAGCAGCGGGCGCTGTGGACTGTGCTTGTGTC 127602
 QY 777 PheLeuArgLeuLeuLeu----- 782
 Db 127601 TTCAATCCGCTGCTGCTGTGTGTGTCTCCGCGCGCCCTCGGCGCTCCCTTCACCTCC 127542
 QY 782 ----- 782
 Db 127541 CTGTCTCTGGCCAGGCTCTCCCTCACCCACTCGCGGAGAGACTGGAGACTCTGCAG 127482
 QY 782 ----- 782
 Db 127481 CCAGGGCCCAAGGCGCACCTGCTCCCTGGGGACCAGCTCCACCCCTGGAAAGCCCTGTC 127422
 QY 782 ----- 782
 Db 127421 CTGCTGCCCTCAGCTGCTGCTGGGGGGCCAGCTCTTGAAGACCCAGACCCCTTC 127362
 QY 782 ----- 782
 Db 127361 GTGCTGTCTCTGCGGTGCTGCTGTGTACCTCGAGCCCTCATCATCTGGGACATATGC 127302
 QY 782 ----- 782
 Db 127301 CCCAGACAGCATGTCCACACCTGACAGGCGCATGCCGACCCCTTGGGCTGTGACGC 127242
 QY 782 ----- 782
 Db 127241 CCCAGTGGTGGCAGAGGGGGCTGCTGGCTGTGGGGTCCATGCCCTAGGCTGTGGAG 127182
 QY 782 ----- 782
 Db 127181 CAGATGGAGCAGAGCTCTGTTTTGGCCCTGCGGGGAGAGACGCTGAGGGTGGCCCT 127122
 QY 783 -----AsnTyrArgIysSerSerArgHisLeuAlaLeuPheLeuAsnLysPhe 798
 Db 127121 TTTGCTTTGCAGAAATTAAGAGAGTCTCCCGCCATCTGGCTGCTTATCAACAGATT 127062
 QY 799 ValGlnPheLeuHisLysTyrTleThrTyrAsnAlaProAlaAlaLeuSerPheLeuGln 818
 Db 127061 GTGCAAGTTCATTCATTAAGTACATTACATACATGCCCCAGCAGCATCTCTCTGAGAG 127002
 QY 819 LysHisAlaAspProLeuHis----- 825
 Db 127001 AAGCAGCCGACCCGCTCCAGTGAAGTCCACAGGGCTCAGGGAAGCTCGGGCTCCGG 126942
 QY 825 ----- 825
 Db 126941 CGAGGGTCCAGCGGGGCTGTGAGGGGTGTTCTTGGGAAGCCGTGTGTGATATGCT 126882
 QY 825 ----- 825
 Db 126881 GCTCCTGAGGCCGGGAGAGGGCTGTGAGCGTGTGTCTATCCAGAGGCGCAGAGCCGCC 126822
 QY 825 ----- 825
 Db 126821 CCACCTGTGCCGAAGCCGCTGTAACCCACCCCTGACCTCTGTCTGTCCGCTT 126762
 QY 825 ----- 825
 Db 126761 CCAGCAGCATGTGGCTTCCCGGCCCGGCGCTTCCCTGTAGAGGTCTGGGCTCTCTGCAT 126702
 QY 825 ----- 825
 Db 126701 GCCTGTTCACACAGCTCCCATGCCACCCCTAGTTGACCCCTGACCCAGCTTCTCTG 126642
 QY 826 -----AspLeuSerPheAspAsn 831
 Db 126641 CCAGCGTCTTCTCCAGGTGAAGACATCTTCTCTCCAGGAGCAGCTGCTTCGACAAAC 126582
 QY 832 SerAspLeuValMetLeuLysSerLeuLeuHisGlyLeuSerLeuProSerArgAspAsn 851
 Db 126581 AGTGACCTGTGTATGTCTAAATCCCTCTGTGAGGGCTGACGCTGCCAGCAGGAGGACAC 126522
 QY 852 ArgThrAspArgGlyLeuAspGluGlu----- 862
 Db 126521 AAGACCGACCGAGGCGCTGTGACGAAGAGGGCGAGGGTGAAGTGGGGGCTGCCAGGGCGCTGCTG 126462
 QY 862 ----- 862
 Db 126461 GAAGCAGCTGGGGAACAGTCTCTACAGGCGCTGTTTTCCAGCGCAGACCGCCCTGGG 126402
 QY 862 ----- 862
 Db 126401 CCGTCCGAGGCGTCCCGATGGGGGCTGTGTAGCCTGGGACCTTGTGTGCCCCCTGA 126342
 QY 862 ----- 862
 Db 126341 CCCACCTCAAGGGGAGCAAGACAGTGAAGTGTGCTGACAGAGACATCAGACACATGCCCC 126282
 QY 862 ----- 862
 Db 126281 CCGGCACATCGCCCGCCCGGCAATGACCCTCCACACCCCGCTCCCTGGACATCC 126222
 QY 862 ----- 862
 Db 126221 ACCATCCACAGTGTGTGTGCGAGGTGTCTCCAGAGAGCTCCCTGTGTCTACTG 126162
 QY 862 ----- 862
 Db 126161 GGCAGAGCTTGTGTTTTTCCCTGCTGAGGTGACACGCGACCGACAGCTGTCT 126102
 QY 862 ----- 862
 Db 126101 CCTGTGCCCGGAGAGAGCGCCAGGCGCGGCTCCCAAGGGGCCATCCGCTTGTGCCCA 126042
 QY 862 ----- 862
 Db 126041 GTGGGTCTGGGCCACATTTGCGAAGAACCCAGCAGGAGAGACTGGCAGAGGTGCATGCG 125982
 QY 862 ----- 862
 Db 125981 CCGATCCAAAGGAGAGCTCCGGGCTCTGTGCTCACCCCATTTGTGAGCGCTGCTGTGT 125922
 QY 862 ----- 862
 Db 125921 GCAGGCACTTGGAGAGCTGGGCAAGTGGCCGCGCTGTCTGCGGAGGCGAGGTCTGTGGT 125862

QY 863 -----G1 863
 Db 125861 CCTCAGTGGCAGCTGCTGGCCGCCCTCCTGCAGCAGCTGCCACTGCTCTCCAGA 125802
 QY 863 ugiuserSerAlaGlySerLeuProLeuValSerValSerLeuPheThrProLeuThrAl 883
 Db 125801 GGAGAGCTCAGCGCGGCTCTGGCCCTGGTGCAGGCTGCTCCCTGTCAACCCCTCTACCCG 125742
 QY 883 aAlaGluMetAlaProTyrMetLysArgLeuSerArgGlyGlnThrValGlu 900
 Db 125741 GGCCGAGATGGCCCCCTACATGAACGGCTTTCGCCGGGCCAAACGGGTGAGGGTGAAGTC 125682
 QY 900 ----- 900
 Db 125681 AGGCCCTGCTTCAACCAAGCCAGGTACAGCCCTGGGTCCCATGGGGCTGCGGGCAGCAG 125622
 QY 900 ----- 900
 Db 125621 CCAAGTACCATTTTGGGCACTGACTTTGGCGTGAAGCTGGGCCACAGCTGGGCTGTGGGTGG 125562
 QY 900 ----- 900
 Db 125561 GTGGGCTGTCTCTCAGAGTGGGAGACACCCCGAGAGAGGCGCTTAGAACTCAG3GGGCC 125502
 QY 900 ----- 900
 Db 125501 AGGGAGCTGGAGGTGGCTGCTCTGCTTTCAGGACACAGCAAAAAGCAGGGAGCCCTTAA 125442
 QY 900 ----- 900
 Db 125441 GGTGTGGGGGTACAGCAGAGAAAGTCAAGCGGGTGGGCAAGCACAGCGGTGGGAGT 125382
 QY 900 ----- 900
 Db 125381 ACCGGTGGCCGCTGTCAAGAGACAGCCAGGCTGAGCGGCCGAGTAGACTCG3CCGTG 125322
 QY 900 ----- 900
 Db 125321 GGGACACTTGTGCTAAGACACCCTGGATATGGGGCTGCCGTGAGGGGTCAAC3CCTCT 125262
 QY 900 ----- 900
 Db 125261 CCTTGGAGCCACTAGTAGGGGAGGCGCGGATTTGTGAGCTTGGGGCAG3TGACT 125202
 QY 900 ----- 900
 Db 125201 GCGTGGGGGCTCACTGGGCTCATGCTTTTCCAGTGTCTTGTGCTC3GGGCC 125142
 QY 900 ----- 900
 Db 125141 TTGGCCCCCTCTGGGTGGCTGGGGTGGGAGAGGTCTGCCCTGGCCCTTACGTTCT 125082
 QY 901 -----AspleuLeuGluValLeuSerAspIleAspIleMetS 913
 Db 125081 GGTCTGGCCCTCTCTCTGTCTGATCATCTGCTGAGGTTCTGAGACATAGACAGATGT 125022
 QY 913 eArpArgArgArgProGluIleLeuSerPhePhe 923
 Db 125021 CCGGGGAGAGACCCGAGATCTTCTTCTGCGGTGAGCTGAGTGGGGCTGAGGGAGA 124962
 QY 923 ----- 923
 Db 124961 GCTGGAGAGACATCTCCCTTGGGGAACCCAGACAGGCGCACCAACCATATGGCTGC 124902
 QY 923 ----- 923
 Db 124901 CCGCGTGAAGCCAGAGAGATCTCTGAGGCTTGGGGTGGGGGCCAGGCCAGG 124842
 QY 923 ----- 923
 Db 124841 TGTTCACAGTCAAGCGCGCCAGAGAGGCTGCTTACGCCCGGGGCTGGAGCCGCGCC 124782

QY 923 ----- 923
 Db 124781 AGTGGACGTGCTCTGCTGCTTGTGATAGGGCCCTAGGCTCCCTGCCGCTTGCAGCCAAC 124722
 QY 923 ----- 923
 Db 124721 CATCAGAGACCAAGAGGTGAAGGGGCTTGTATGGGGGCTGAGACCCCTTGTCAACACC 124662
 QY 924 -----Ser-ThrAsnLeuGlnArgLeuMetSerSerAlaGluGlyCys 937
 Db 124661 TTCCACCTTTCCCGGCTGACAGCAACCTGCACCGGCTGTATGAGCTCGGCCGAGAGTGT 124602
 QY 938 CysArgAsnLeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerPro 954
 Db 124601 TGGCGACACTGGCTTCAAGCTTGGCCCTGCGCTGCTCCATGAGAACAGCCCAAGGTAGGCC 124542
 QY 954 ----- 954
 Db 124541 CACGGCGGCTCTGTGCCCGCGCACCGCTGGGGCCCTGTAAACCAAGTCAACCGCTTGA 124482
 QY 954 ----- 954
 Db 124481 CCGCGCTGTGGCCAGACAGGCTTCCAGCGGCTCAGTCGCGCTGTGGGGGCTC 124422
 QY 955 -----SerIleAlaAla 958
 Db 124421 TGGCGTGGGACAGAGGTGAGGTTCCCACTGAAGTGTGGGCTTTCAACATTGCAAGCC 124362
 QY 959 AlaPheLeuProThrPheMetTyrCysLeuGlySerGlnAspPheGluValGlnThr 978
 Db 124361 GCTTCTCTGCCACAGTTCATGATCTGCTGGGACAGCAGACTTTGTGAGTGTGCAGACG 124302
 QY 979 AlaLeuArgAsnLeuProGluTyrAlaLeuLeuCysGln 991
 Db 124301 GCCCTCGGAACCTGCTGATGAGCTCTCTGTGCCAAGTGAAGACACTGGCCCCGCC 124242
 QY 991 ----- 991
 Db 124241 CAGGGCCCTGGCAAGCCCTGGGGGAGGCGGAGGTGGCTCCCGAGAGGCGTGGCTTGGG 124182
 QY 991 ----- 991
 Db 124181 AGTCCGAGGCTGTGCTACTGCGGCGCAGAGTGTGTGGCCAGGCTGACGTGAGGGCCCC 124122
 QY 992 -----GluHisAlaIleValLeuLeuHisArgAlaPheLeuValGlyMetTyr 1007
 Db 124121 GTTTCGCCACAGAGACAGCGGCTGTGCTGCTCACCGGGCTTCTGTTGGGCAATGTAC 124062
 QY 1008 GlyIleMetAspProSerAlaGlnIleSerGluAlaLeuArgIleLeuHisMetGluAla 1027
 Db 124061 GGCCAGATGAGACCCAGCGGCGAGATCTCGAGGGCCCTGAGATTCGCAATATGAGAGCC 124002
 QY 1028 ValMet 1029
 Db 124001 GTGATG 123996
 RESULT 9
 AC130221
 LOCUS AC130221 219964 bp DNA linear HTG 08-AUG-2002
 DEFINITION Mus musculus chromosome UNK clone RP23-168B11, WORKING DRAFT
 AC130221
 ACCESSION AC130221
 VERSION AC130221.1 GI:22138681
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 219964)
 AUTHORS McPherson,J.D. and Waterson,R.H.
 TITLE The sequence of Mus musculus clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 219964)

AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-AUG-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@wustl.wustl.edu
 Project Information -----
 Center project name: M_BA0168E11

----- Summary Statistics -----
 Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-terminator; 0% of reads
 Chemistry: Dye-terminator; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 217436 bases at least Q40
 Consensus quality: 217966 bases at least Q30
 Consensus quality: 218400 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 2186: contig of 2186 bp in length
 * 2187 2286: gap of unknown length
 * 2287 6318: contig of 4032 bp in length
 * 6319 6418: gap of unknown length
 * 6419 33825: contig of 27407 bp in length
 * 33826 33925: gap of unknown length
 * 33926 80048: contig of 46123 bp in length
 * 80049 80148: gap of unknown length
 * 80149 128323: contig of 48175 bp in length
 * 128324 128423: gap of unknown length
 * 128424 219964: contig of 91541 bp in length.

FEATURES
 source
 Location/Qualifiers
 1. 219964

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="YUNK"

/clone="RP23-168E11"

misc_feature

/note="assembly_name:Contig13"

misc_feature

/note="assembly_name:Contig14"

misc_feature

/note="assembly_name:Contig15"

misc_feature

/note="assembly_name:Contig16"

misc_feature

/note="assembly_name:Contig17"

misc_feature

/note="assembly_name:Contig18"

BASE COUNT 51706 a 56458 c 58030 g 53268 t 502 others

ORIGIN

Alignment Scores:

Pred. No.: 8 96e-154 Length: 219964
 Score: 2673.50 Matches: 844
 Percent Similarity: 27.70% Conservative: 75
 Best Local Similarity: 25.44% Mismatches: 93
 Query Match: 51.01% Indels: 2317
 DB: 2 Gaps: 22

US-09-929-769-7 (1-1029) x AC130221 (1-219964)

QY 18 PROATGAlaAspSperSerGluPheGlnAlaLeuLeuAspPleThrPheProGluLys 37
 |||||
 Db 155249 CCTAAGCTGGTACAGAGGAGTGTGAGTGTGCTGGACATCTGGTTCCAGAGAAG 155308
 QY 38 ProLeuProThrAlaPheLeuValAspThrSerGluGluAlaLeuLeuProAspThr 57
 |||||
 Db 155309 CCTTGGCCACAGCCCTTCTGTGTGACACGTGGAGAGGCGCTGTCTGCTACGTGG 155368
 QY 58 LeuLysLeuArgMetIleArgSerGluValLeuArgLeuValAspAla 73
 |||||
 Db 155369 CTGAAGCTGGCAATGATGCTGTGAGGTCCCGCTGTGGAGCC-TGGTACAGTGA 155427
 QY 73 73
 Db 155428 GGTAGGGAGGGGCTCGGCTCAGCTGCTGCCGGGGAGGCTAACTGACTGAGTCA 155487
 QY 73 73
 Db 155488 CCTTCCAGGACTGTGAGCGCTCAGAGACAGGCTGAGAGAAAGGCTATAGAGACC 155547
 QY 73 73
 Db 155548 TGGCTTTAGAGGCACAAAGATAGTTCAGAGGCTTTGGAGAACATGTGGCTATAGGC 155607
 QY 73 73
 Db 155608 AAGCTGCACTGACAGAGGAAGTCCATCGGGGTAGGCGTGGCCAGCTTGGGCTTT 155667
 QY 73 73
 Db 155668 CAGTCTGGCCCGCAGCAGGAAAGAGACATCCATGCTGTGGTGTGAGCTGGGGAAGCC 155727
 QY 73 73
 Db 155728 AGTGAAGACATACCATGAGTCTGTGACTGTAGTGGCCAGAGGTCCATTTGCAATGTCA 155787
 QY 73 73
 Db 155788 TTGCTAAAGAACCATGGGGGAGTACTATTTGGCAGAGCCAGGCTTGTGCTGTGACCC 155847
 QY 74 74
 Db 155848 TGAGATACCTTTCAGAGCAGCAGGCGCCAGCTGAGCATCTTTCACACACCCCTGCAG 155907
 QY 77 AspleuGluProGlnGlnLeuLeuPheValGlnSerPheGlyIleProValSerSer 96
 |||||
 Db 155908 GACCTGGAGCCCGCAGCAGTGTCTCTTCTGTGAGTCTTGTGGCATCCCGTGTCCAGC 155967
 QY 97 MetSerLysLeuLeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGlu 116
 |||||
 Db 155968 ATGACCAAGCTGTCTGAGTACTGTGACAGCGGTAGGCCAGGACCCCAAAACCTGGAG 156027
 QY 117 GlnAsnIleMetAspLys 122
 |||||
 Db 156028 CAGAACATCATGGACAAAGTAAGCCCTGGGGCGCGAGCTGGCAGACACCTGAGGCC 156087
 QY 122 122
 Db 156088 CACCTAACCCCTTAGGGCAGATGTGTGCTGTGGGTGAGTGGGACACTACCTGTG 156147
 QY 122 122
 Db 156148 GCCAAGCTCCAGAGCTGTGAGACTGACTGTGCATCACTACTGAGCTGAGTGAAGA 156207
 QY 122 122
 Db 156208 GACCCAGCAGCTCAGTCCCTGAGCTTTGTTTGTGTTTGTGACCTTTTGAGACAGTGT 156267
 QY 122 122
 Db 156268 TCTCTGTAGCTTGGCTATCTCTGAACCTAGCTTTGTAGACCAAGCTAGCTCAACTC 156327
 QY 122 122

Dh	156328	AGAGATCTTGCCCTGCTGCTCTGCTTAAAGTGTACAGACTTAAAGTGTGTGTGTCCACACACTGGCCA	156387
Qy	122	-----	122
Dh	156388	GCTAGCTCTTCGACAGGCATAGAGAGACTCGGTCATGTGGGTGTGTGCCAGCCCCCCCC	156447
Qy	122	-----	122
Dh	156448	AGTCTCATGTTTGGTGGAAAAAGCAGGCTGGCTTCACCCACCCCAAGTGTCCCTGGGTG	156507
Qy	123	-----AsnThrMetAlaHisLeuValGluValAlaGlnHisGluArgValAlaSerGlu	139
Dh	156508	CTGTTTTCGACATATTACATATGGCCCCACCTGTGTGAGGTACAGACAGAGAGAGTGTCTGTG	156567
Qy	139	yGlyGlnThrPheHisSerLeuLeuThrAlaSerLeuProProArgArg-----	155
Dh	156568	AGGCAGACCTTCCACTACACTCCACAGAGCTCCCTCCACACCCCGAGAGGTACAGCCCC	156627
Qy	155	-----	155
Dh	156628	CTCCATTACCCCTACATCTTACTCCCGAGCTGGACCGGCTTTTGGCTCCATGTGTTC	156687
Qy	155	-----	155
Dh	156688	CTGGCTTGTCCACAGCTTATTTTGGGGGCTCTTATGTGGGCCAAGGTCAGACCTGTCTT	156747
Qy	155	-----	155
Dh	156748	CTGGTCACAGTCGCTGGCGAGGACAGACTGTGGCAGAGCTGTGTGCCAAGTTTGAAGCC	156807
Qy	156	-----AspSerThrGluAlaProLysProLysSerSerProGlu	169
Dh	156808	ACACTTCTCTGTTCTTCGACAGACAGCAGGAGGCTCTTAAGCCAGAAGCAGCCCTBAGC	156867
Qy	169	LeuProIleGlyGlnGlyArgGlyLeuArgValGlyThrGlnLeuArgValLeuGlyProGlu	189
Dh	156868	CCCCACACGAGCCAGGAGGAGGAGCTGGCGAGGAGGAGCGAGGTGCCGCTTGGCCCTBAGG	156927
Qy	189	spAspLeuAlaGlyMetPheLeu-----	196
Dh	156928	ACGACTTAGCTGGCATCTTCTCT-ACAGGTACAAAGGCTACCGTCAGAGGGCCCCCTACGA	156988
Qy	196	-----	196
Dh	156987	GCACCGCCTGCACATGTGTGCCCTGCATGTGACTCGTACATGTGTGTCTGGCCACTATT	157048
Qy	196	-----	196
Dh	157047	CCATTCTTATCTTGTGTCTCACTCTTCATGTCACTGTATGTAAGTGTGCTGTGGACCCC	157106
Qy	196	-----	196
Dh	157107	TGTGTGTCCCTGTGTGTAGGGGTGTAGGTACTGTGTCTTCCTTCATGGCAGACTCTTG3GGTGT	157166
Qy	196	-----	196
Dh	157167	TAGACTCCCTGTGCCACTCATCAGCAAAAGACTGTGGGTCCCTTAGTTAATGACACTGGGA	157226
Qy	196	-----	196
Dh	157227	CGGAATCTGATGGCGCTGAGCAGGCTACGCAATGGAGCCACGGGTGCGGGGCTCTGCG	157286
Qy	197	-----GlnIlePheProLeuSerProAspProArgTyrIleProIns	210
Dh	157287	TTCCTTATCTTCTGTCTCTCCAGATCTTCCACTGAGCCCGGAGCCACGGTGGCAGAGCT	157348
Qy	210	erSerProArgProValAlaLeuAlaLeuGlnGlnAlaLeuGlyGlnGluLeuAlaArg	230
Dh	157347	CGAGCCACAGGCCCCCTTGGCCTGGCAGTCGACGAAGCCCTGTGGCCAAAGAGTGTGCTGTG	157406
Qy	230	aValGlnGlySerProGluValProGlyIleThrValArgValLeuGlnAlaLeuAlaT	250

[illegible]

QY 363 a11leAlaThrLeuArgSerGlyIuGInCysSerValGluProAspLeuIleSerLysV 383
 |||||
 Db 158546 TCATTGCCACCTGTAGAGTCGGGGAGAAAGTGTCTGTGAGCGCTGACATCAGCAAG 158605
 QY 383 a1----- 383
 Db 158606 G-TACTGTCCCTCCCTCTGCCCCACGGAGAGTCTCGGAGAGATGGGGAGTGTCTGC 158664
 QY 383 ----- 383
 Db 158665 CTCAGGGAACACAGAGCCCTTGCCCTCTCCTCAGCCAGCTCCTCTGTGTCCATAG 158724
 QY 384 --LeuGInGlyLeuIleGluValArgSerProHisLeuGInIleLeuLeuThrAlaPhe 403
 |||||
 Db 158725 TCCTCGGGGGCTCATTCAGAGTGGCGTCTCCCGCAGCTGGAAGATGTCGACAGCCCTCT 158784
 QY 403 heSerAlaThrAlaAspAlaIleSerProPheProAlaCysLysProValValValS 423
 |||||
 Db 158785 TCCTAGCCACACAGAGACCTGTGCGCCAGCCCTCTGCGCCATTTGTGTGTGA 158844
 QY 423 erSerLeuLeuGInGInGInIuGInProLeuAlaGlyLysProGlyAlaAspGlyG 443
 |||||
 Db 158845 GCTCCTGTCTGCTACAGAGAAAGAGAGAGCTCTCGGGCC-CAGCAAGCAGAGAGGTGAG 158903
 QY 443 ly----- 443
 Db 158904 GGTGCCAGTAAGTGGAGATGTGGCGGTGGGCCAGGGGGGTGCTTCAGACCTAGTC 158963
 QY 444 -----SerLeuGInAlaValAlaArgLeuGlyProSerSerGlyL 456
 |||||
 Db 158964 TCCTGCTGACCCCTGTCTGCCAGACGAGACCGATGCGCTGCGGACAGCTTCGCGGC 159023
 QY 456 euLeuValAspTrpLeuGInuMetLeuAspProGluValAlaSerSerCysProAspLeuG 476
 |||||
 Db 159024 TCCTGTGACTGTGCTGGAGAGCTGGACCCGAAAGTGTGTGTACTGCTCCCTGACTAC 159083
 QY 476 InLeuAlaArgLeuLeuPheSerArgArgLys----- 485
 |||||
 Db 159084 ACTGGAAGCTGCTCTCTCAGAGAAAGGTGACAGAGACATGGCCGGGTCTCTCCAC 159143
 QY 485 ----- 485
 Db 159144 AGTGTGCCACTCAGACTATGTCTGTAGCCTCAGAGGGGTGGGTTAGCCAGGCTGTG 159203
 QY 485 ----- 485
 Db 159204 GGCTCCCTGTCCAGAGAGTGAAGTGTGCATGTCTTTGGGGCTGGAATGGGCTAGGC 159263
 QY 485 ----- 485
 Db 159264 AGGATAGTCCACAGAGCCTAAAGTCTGGGAAACCAGGTAGGCCAGGAGATGTGGGAA 159323
 QY 485 ----- 485
 Db 159324 TGAGGGAGGGAGTGTGGCTGTGGCTGTCTAATTAACCCAGCCCTTGTCCAGGGCAGAGA 159383
 QY 485 ----- 485
 Db 159384 GAAGATGGCCAGAGCCCTTAGCTGTCTCATGGTGGCCCTACCAGAACATGTCACTGA 159443
 QY 485 ----- 485
 Db 159444 CATGCTGTCTCAGAACTGTGGCCCTTGCCCAATTGAGATCACCCTTGCCCTTTTCCTC 159503
 QY 486 -----GlyLysGly-----GlnAlaGlnValProSerPheArgProTyrLeuLeuThrL 502
 |||||
 Db 159504 TCACAGGCAAGGACCATTAAGTGCAGGTCTCTCTTCGCGCCCTCCTCCTGCGCC 159563
 QY 502 euPheThrHisGlnSerSerTrpProThrLeuHisGlnCysIleArgValLeuLeuGlyL 522
 |||||
 Db 159564 TCCTACGACGACAGGCGAGCTGTGCACGCTGACTGCTGATCCGTGTCTGTCTAGGCA 159623

QY 522 ysSerArgIuGIn----- 526
 |||||
 Db 159624 AGAGCGGTGAGCAGAGAGTGAAGCCCACTGACAGACCCAGAGAGCTCTCGGCTGATGCA 159683
 QY 526 ----- 526
 Db 159684 ACTTTGTGTAAACATGAATTTGTTCCTTCCACTCTATTCAGGGCAGCATCTCTGAG 159743
 QY 526 ----- 526
 Db 159744 GGGGGTGTATTGTGTCTCAGTGTCTTGGCGACTGCAAGCCCTCAATAGCAGGGCCAGTA 159803
 QY 526 ----- 526
 Db 159804 TTTTGTTTTGTGGCCAGCCCTCAGCTAGGGGTAGACAGTCCAGGCCCTCCATCTCC 159863
 QY 526 ----- 526
 Db 159864 TTGGCACAATTTTGTCTGTGAGAAATTTGGCTTACGTCCAGCCCATGTACAGAGGCAA 159923
 QY 526 ----- 526
 Db 159924 AAGAGCTGTGGCAGAGCTGCCCATCAGACCTGTGTCTTGTACTGTGACAGGCCATAGT 159983
 QY 526 ----- 526
 Db 159984 GTGGGACAGAGCTGTGCTGTGATTTTCTCATAGGAACCATCAGTCCACAGTCACAAAC 160043
 QY 526 ----- 526
 Db 160044 ACAAGACTGTGGAGTGTGTGGGTGACAGGAGACCAAGACTAGGAGACTCTCAGAGAGACA 160103
 QY 526 ----- 526
 Db 160104 GAGGTGGCCCTCTACTGAGGTCCCTTTTATAGAACTTTACCTGGCATTTGCTGCTCT 160163
 QY 527 --ArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCysIleHisValProArg 545
 |||||
 Db 160164 CCAGAGCTGACCCCTTCAGATCCCTGAGACTTCTGTGGCTGTATCCATGTCCCCCG 160223
 QY 546 IleTrpGInGlyArgAspGlnArgThrPro----- 555
 |||||
 Db 160224 ATCTGGAGGGCCGGAGACGCGCACTCCACAGGAGGTCTCAAGAGGTTAGCCCAACC 160283
 QY 556 -----GlnLysArg 558
 |||||
 Db 160284 ATGGGTGTGGCTGCACACATCTGGGGCTACCTGCAGCCTCCTGATTCAGAGAGCG 160343
 QY 558 gatrgIuGInLeuValLeuArgValGInGlyProGluLeuIleSerLeuValGluLeuI 578
 |||||
 Db 160344 GCGGAGAGAGTGTGTCTCCAGCTGCAGGGCCAGAGCTGTCTCAGCCTGTGTGAGCTTAT 160403
 QY 578 eleuAlaGluAlaGlnThrArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIleG 598
 |||||
 Db 160404 CCTGTCCGAGGACGAGACAGAGGAGATGGGACAGTGTGTCAGCGACCTTATCCA 160463
 QY 598 nAlaArgLeuProLeuLeuLeuSerCysCysGlyAspAspGlnSerValArgLysVal 618
 |||||
 Db 160464 GACCGGCTGCCCTCTCTCTTACTGTCTGTCTCCCTAGCAATGATGAGCATTTGGAAAGT 160523
 QY 618 lThrGlnHisLeuSerGlyCysIleGInGInTrpGlyAspSer----- 632
 |||||
 Db 160524 GACCGAGACCTGACAGCTGCATTCAGCAATGGGGGAGAG-GTACGTGAGTCCCACT 160582
 QY 632 ----- 632
 Db 160583 GTGTCCCTTGCCATAGATGCTGTGTCTACCTGGAGTGGCTCTTCAGTACATGTC 160642
 QY 633 -----Val 633
 |||||
 Db 160643 CCTCCCTGGCCACACCATCCAGTACCCAGTCCCTCTCTCTCTCTTCCAGTGT 160702
 QY 633 lLeuGlyArgArgCysArgAspLeuLeuGInLeuTyrLeuGlnArgProGluLeuArg 653

Db 160703	CGTGGGCCACGGCGTCCCGACAGCTTGGCTGTGTGCACCTGACTGTCGAGGGCCAGACGGTCCG	160762
QY	653 gValProValProGluValLeuLeuHisSerGluGlyAlaLaserSerSerValCysIly	673
Db 160763	GCTGCCAGTGGCCCGGAGGTCTCGTGTGCAAAAGTGAAGGTGGCCACAGCAGCAGCTTGTGCA	160822
QY	673 sleu-----	674
Db 160823	GCT-GAGGGACACAGTCTCCACAGACGAGGGGGCTGGGGGCTGCAGAGATGGGGTGGGGG	160881
QY	674 -----	674
Db 160882	GGGGGTGTCACCAATCAACCTGCACCAAGAAGTGAAGGGGCTATGTCCTATGAGCCTGGAG	160941
QY	674 -----	674
Db 160942	TGGGGGTGCATCAAGTGTCACTTAACTCACTCACTCCAGAGCTAGACCTTACTGTCTAG	161001
QY	674 -----	674
Db 161002	GCTGTCTCTGTGAGTGCAGCTGTGGGGGCGAGAGGATGGCCCTTCCCTTTGTGTACCCAGG	161061
QY	674 -----	674
Db 161062	TGTGGGCGACAGAGGAGTCAACACATGTGTGTCTATGTGACAAATATGGGGCATGCGCTTC	161121
QY	674 -----	674
Db 161122	CATCAGAGGGGTGCTGCACCTGTCAACCCCAAGTACCCCAAGGTCTCCACCTCTTACCCA	161181
QY	674 -----	674
Db 161182	TGTTCGGCAGCTCCACTGTGGCTGCAGAGGGGACAGCAGCTTGGCGTCTGGAGATTGG	161241
QY	674 -----	674
Db 161242	TGTGAGGGGCCACATGAGCTCTTGCCTTGTGTGATGCTGACAGTCTCGGTGCTGAGCC	161301
QY	674 -----	674
Db 161302	TGTGTCTGTCTGTGCTGATGTCTGTGGGCCATGAGGGGGGTAGGGGGAGTTCACAGCGGTG	161361
QY	675 -----	675
Db 161362	ATGAGCAAGTATGTCCCCCTCCCTTACGTGGAGCGGCTGTGCATTCATCAGCT	161421
QY	684 uLeuAlaSprHisSerAspSerArgAlaLeuGluAsnArgGlyAlaAspAlaSerMetAl	704
Db 161422	CTGTGCAGACAGCAGCGACTCCCGGTGTGTGTGAGAGAGAGAGTGTGAGCCCAACATGGC	161481
QY	704 aCysArgGlySerLeuAlaValAlaHisProLeuLeuLeu-----	717
Db 161482	CTGCCGGAAGCTGTGGCTGTGGCCACCCCGCTGTCTGTCT-GAGGTGCTGTCTCACAGC	161540
QY	718 -----	718
Db 161541	CTGTGTCTGTCTGGGTGACTCGGGGGGAAGGTGTAAACCCAGAGCAATTTCTTCTCTGCCCA	161600
QY	718 gHisLeuProMetIleAlaAlaLeuLeuHisGlyAlaArgThrHisLeuAsnPhelGlnIupH	738
Db 161601	GCACCTGCCCAATGTATCGGAGGCGCTCTCGACAGCGCCGACATCATCTGAACTCCAGAGATT	161660
QY	738 aArgGlnGlnAsnHisLeuSerCysPheLeuHisValLeuGluLeuLeuLeuLeuLeuG	758
Db 161661	CCGGCAGACAGAACCACTGGGCTTTCTTCTGTGCATGTCTGGGATCTCTGGAGCTGTGCA	161720
QY	758 nProHisValPheArgSerGluHisGlnGlyAlaLeuThrAspCysLeuLeuSerPheI	778
Db 161721	GCCTGGTGTGTTCACAGAGCCAAACSCAGGGGGGCGCTTGGGAGCTGCTGGTGCATTCCTTAT	161780
QY	778 eArgLeuLeuLeu-----	782

[illegible]


```

|||||
Db 165078 GTCCTGTCAGACAGACCTTCCTTGTGGCGGTACGCGCCAGATTGACACAGCGCCAG 165137
OY 1016 IISerGluAlaLeuArgIleLeuHisMetGluAlaValMet 1029
Db 165138 ATATCAGAGCCCTGGAAGATCTTACACATGAGAGCGCTGTGATG 165179

RESULT 10
HSM802043 1141 bp mRNA linear PRI 18-FEB-2000
LOCUS Homo sapiens mRNA: cDNA DKFp434C0126. (from clone DKFp434C0126);
DEFINITION partial cds.
ACCESSION AL137358
VERSION AL137358.1 GI:6807876
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1141)
AUTHORS Ansoorge, W., Winkner, U., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFp434C0126) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
Location/Qualifiers
source 1. 1141
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFp434C0126"
/cisue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1. 934
/gene="DKFp434C0126"
<1. 934
/gene="DKFp434C0126"
/note="DKFp434C0126, partial"
/codon_start=2
/product="hypothetical protein"
/protein_id="CA870710.1"
/db_xref="GI:6807877"
/translacion="LIQPHVFRSEHOGALMDCLIFRLLINRKSSRLIAFINKEV
OFIRKITYNAPALISFLQKADPDLSEFNDDYMLSLAVLSLPSDDDEDTGL
DEGEESGASGSLPIVSVLFPITAAKAPYKRLSRQTVYDLDLEVLSDIDEMSR
REILSFTSTNQRKMSNAECCRNALFCLARSMNSSTIAAEFLPTMYCIGSODF
EYVQALRNLPETALCOEHAVALHRAFGVHVRDGFQADLRKPEDPAYGGRDVS
LWQPTPLQAPARVPDPRKQAEAWALLVCPRR"
polya_site 1126
BASE COUNT 209 a 384 c 339 g 209 t
ORIGIN
Alignment Scores:
Pred. No.: 1.69e-75 Length: 1141
Score: 1359.00 Matches: 272
Percent Similarity: 98.91% Conservative: 0
Best Local Similarity: 98.91% Mismatches: 2
Query Match: 25.93% Indels: 1
DB: 9 Gaps: 0
US-09-929-769-7 (1-1029) x HSM802043 (1-1141)
OY 756 LeuLeuGlnProHisValPheArgSerGluHisGlnGlnValAlaLeuTrpAspCysLeuLeu 775

```

```

|||||
Db 2 CTGCTGACGCGCAGCAGCTGTCCGACGACAGCAGGCGCGCTGTGGACTGCTCTCTG 61
OY 776 SerPheIleArgLeuLeuAsnTrpArgLysSerSerArgHisLeuAlaIlePheIle 795
Db 62 TCCCTTATCCGCGCTGCTGCTGAATACAGAAAGCTCCCGCATCTGGCTGCTTATC 121
OY 796 AsnLysPheValGlnPheIleHisLysTrpIleThrTrpAsnAlaProAlaAlaIleSer 815
Db 122 AACAGATTGTGAGAGTTCATTCATTAAGTACATTACATGAGCCAGCAGCATCTCC 181
OY 816 PheLeuGlnLysHisAlaAspProLeuHisAspLeuSerPheAspAsnSerAspLeuVal 835
Db 182 TTCCTGACAGAGCAGCGCGCAGCTCCAGACCTGCTTCGCAACACTGACTGCTGTG 241
OY 836 MetLeuLysSerLeuLeuAlaGlyLeuSerLeuProSerArgAspAspArgTrpAspArg 855
Db 242 ATGCTGAATTCCTCTCTTGCAGTCTCAGCTGCCACGACGACAGCAGCAGCAGCA 301
OY 856 GlyLeuAspGluGluGluGluGluGluSerSerAlaGlySerLeuProLeuValSerVal 875
Db 302 GGCCTGACAGAGAGGCGGAGAGAGAGAGCTCAGCGCGCTCCCTGGCCCTGTGAGGCTC 361
OY 876 SerLeuPheThrProLeuThrAlaAlaGluMetAlaProTrpMetLysArgLeuSerArg 895
Db 362 TCCCTGTTCACCCCTCTGACCGCGCGGAGATGCCCCCTACATGAACGCGCTTCCCGG 421
OY 896 GlyGlnThrValGluAspLeuLeuGluValLeuSerAspIleAspGluMetSerArgArg 915
Db 422 GCCCAACGCTGACAGATCTGCTGAGAGTCTGAGTACATACAGAGATGCTCCGCGCG 481
OY 916 ArgProGluIleLeuSerPhePheSerThrAsnLeuGlnArgLeuMetSerSerAlaGlu 935
Db 482 AGACCCGAGATCTCTGACCTTCTCTGACCAACCTGACGCGGTGATGAGACTGCGCGAG 541
OY 936 GluGlyCysArgAsnLeuAlaPheSerLeuAlaLeuArgSerMetGlnAsnSerProSer 955
Db 542 GAGGTGTCGCGCAGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
OY 956 IleAlaIleAlaPheLeuProThrPheMetTrpCysLeuGluSerGlnAspPheGluVal 975
Db 602 ATTCACACGCGCTTCCGCGCAGCTTCATGATGCTGCGGACGACGAGACTTGGAGCTG 661
OY 976 ValGlnThrAlaLeuArgAsnLeuProGluTrpAlaLeuLeuGluGlnGlnHisAlaIle 995
Db 662 GTGACAGAGCGCGCTCCGGAAGCTGCTGAGTACGCTCTCTGCTGCTGCTGCTGCTGCT 721
OY 996 ValLeuLeuHisArgAlaIlePheLeuValGlyMetTrpGlyGlnMetAspProSerAlaG 1015
Db 722 GTGCTGCTCCACGCGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781
OY 1015 IISerGluAlaLeuArgIleLeuHisMetGluAlaValMet 1029
Db 782 GATCTCGAGAGCCCTGAGATCTGATATGAGAGCGCGGTATG 824

RESULT 11
AC020509/c 157851 bp DNA linear HTG 03-JAN-2000
LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
DEFINITION pieces.
ACCESSION AC020509
VERSION AC020509.1 GI:6664388
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 157851)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA.

```


COMMENT

This sequence was identified as CDM:10213488 by the submitter.
For more information on this record e-mail to fly@celera.com.

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

source

1. 157851
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT 41166 a 36910 c 37402 g 42373 t
ORIGIN

Alignment Scores:

Pred. No.:	5.53e-42	Length:	157851
Score:	865.00	Matches:	288
Percent Similarity:	43.408	Conservative:	189
Best Local Similarity:	26.218	Mismatches:	400
Query Match:	16.50%	Indels:	222
		Gaps:	34

US-09-929-769-7 (1-1029) x AC020509 (1-157851)

QY 1 MethisileuValHisalMetValileuLeuThrluLeuGlyProProArgAla 20
Db 82377 CTCACATATAACATCATTCACGCCCTTCATTCCTGCCTCACCTATTCC-----AACAGC 82324
QY 21 AspAspSerGluPheGlnAlaLeuLeuAspIleTrpPheProGluGluLysProLeuPro 40
Db 82323 AACATGCCCGGATCGATACCCATTTTGACTATTGGTTCGCCCGGACACCGACGCCGCC 82264
QY 41 ThrAlaPheLeuValAspThrSerGluGluAlaLeu---LeuLeuProAspTrpLeuLys 59
Db 82263 GTCGCTTTCTGCCACGATCGCCGAGAGACAGATGTCACACTGCTGCCGCTGGAAG 82204
QY 60 LeuArgMetIleArgSerGluValLeuArgLeuValAspAlaAlaLeuGlnAspLeuGlu 79
Db 82203 CTGACATGATCCGCTCATCGGTGAGACAGCTAATTGAGGACGCTCTCAACGATCTAACG 82144
QY 80 ProGlnGlnLeuLeuPheValGlnSerPheGlyIleProValSerSerMetSerLys 99
Db 82143 CCGGATCATCGCTGCTTTTGCAAACTTTGGCAGCGCGGTCAACTCGATGTCGCAAG 82084
QY 100 LeuLeuGlnPheLeuAspGlnAlaValAlaHisAspProGlnThrLeuGlnGlnAsnIle 119
Db 82083 CTCCTGCGCATAGCTGACACCGCTGTACGAGACACTTTGATGTGGTGAAGATGGCATT 82024
QY 120 MetAspLysAsnTrpMetAlaHisLeuValGluValGlnHisGluArgGlyAlaSerGly 139
Db 82023 CTGACAAAGGCGCTATCTGGCCAGCTAATCGAGATTGACAGAGCGCGGTCTAAGAAAT 81964
QY 140 GlnGln---ThrPheHisSerLeuThrlaSerLeuProProArgAspSerThr 158
Db 81963 GGGCACTACACCGCTACAGCGCTTGATCTGCATTCC----- 81928
QY 159 GluAlaProLysProLysSerSerProGluGlnProIleGlyGlnLysArgIleArgVal 178
Db 81927 -----CACTCGCAAGACTGTGCCAGATCTGCC-----AAGACTAGTGTG 81889
QY 179 GlyThrGlnLeuArgVal-----LeuGlyProGluLysAspLeuAla 192
Db 81888 GTTATTACGAGAGCGCTTAATTTGATTGATTACGATTCTTCACAGCTCGATGATAGACC 81829
QY 193 GlyMetPheLeu-----GlnIlePheProLeuSerProAspProArg 206
Db 81828 ACTAAGCTTCTGGCCACCAAGAGGAGTAGCCCAATCTCTCAGCAGCCGCGAC----- 81775
QY 207 TrpGlnSerSerSerProArgProValAlaLeuAlaLeuGlnGlnAlaLeuGlyGlnLys 226
Db 81774 ---CAATTGACGTGAGTGGGAGTGAAGTGCAGCTTGTATTCAAAAACGTGTGCAATG 81718
QY 227 LeuAlaArgValAlaGlnLysSerProGluValProGlyIleThrValArgValLeuGln 246
Db 80952 GGGCTTTCCGAGGTCAAGCTGACAGCACTAATGATTCAAGCTTCTGCTG----- 80899

Db 81717 CTAGCAAGCCCGCAACGATACAGGATGTGGTTAACGCCATTAACAGAGGTCTAGCC 81658
QY 247 AlaLeuAlaThrLeuLeuSerSerProHisGlyValAlaLeuValMetSerMetHisArg 266
Db 81657 GTG-----GTTGACAGTGTACGATGAGCCCTCAC--- 81628
QY 267 SerHisPheLeuAlaCysProLeuLeuArgGlnLeuCysGlnTrpGlnArgCysValPro 286
Db 81627 -----GCCGTCACATTTTAAAG----- 81610
QY 287 GlnAspThrGlyPheSerSerLeuPheLeuLysValLeuLeuGlnMetLeuGlnTrpLeu 306
Db 81609 ---ACTTTCTTCAGCTGATGCTGCACAGCAGCAAGATATCATCTACTAGAG----- 81562
QY 307 AspSerProGlyValGlnGlyGlyProLeuArgAlaGlnLeuArgMetLeuAlaSerGln 326
Db 81561 -----AAGCTCTCCAAAGAACTTAAGTATTT----- 81532
QY 327 AlaSerAlaGlyArgArgLeuSerAspValArgGlyGlyLeuLeuArgLeuAlaGlnAla 346
Db 81531 -----AAGCACATATTCGCCGACATCTGCAGTCTGCGCAAGAAATCCGAA--- 81490
QY 347 LeuAlaPheArgGlnAspLeuGluValValSerSerThrValArgAlaValIleAlaThr 366
Db 81489 ---CTGTATCATGAGAGCTTGCTTTCATGCTGAGAACTCTGCGAGATTTATGCG--- 81436
QY 367 LeuArgSerGlyGlnGlnCysSerValGluProAspLeuLeuSer-----LysValLeu 384
Db 81435 -----CAGCAGTTTAAAGCGAATACCGCTTGTGTGACGAAACGAGATCGTC 81388
QY 385 GlnGlyLeuIleGlyValArgSerProHisLeuGlnGluLeuLeuThrAlaPhePheSer 404
Db 81387 CCGGCGCATCTCCAA-----ACCTTCGAT 81364
QY 405 AlaThrAlaAspAlaAlaSerProPheProAlaCysLysProValValValValSerSer 424
Db 81363 CAGACCAAGCAGACG-----AAGACCGCTGCCCAAGTCCCAAGAGC 81325
QY 425 LeuLeuLeuGlnGlnGlnGluProLeuAlaGlyGlyLysProGlyAlaAspGlyLysSer 444
Db 81324 GACCACCTCTCCAC----- 81310
QY 445 LeuGluAlaValArgLeuGlyProSerSerGlyLeuLeuValAspTrpLeuGluMetLeu 464
Db 81309 -----AAGGGCTCTTCATGCACTGCGCTCGCAATG 81277
QY 465 AspProGluValValSerSerCysProAspLeuGlnLeu-----ArgLeuPhe 481
Db 81276 GATCCGAGATAGTTTCCACT-----CAGTAATGAAAGAGCGCTTCTGTGTC 81229
QY 482 SerArgArgLysGlyLysGlnAlaGlnValProSerPheArgProTyrlaLeuThr 501
Db 81228 TCAGAG-----TCCTGACGAGAGTTTGTTGCTTATCTGTTGTCC 81190
QY 502 LeuPheThrHisGlnSerSerTrpProThrLeuHisGlnLysIleArgValLeuLeuGly 521
Db 81189 CTGATGACACCAACCAACCACTGAGACAGCATGGAAG---ATGGCGGAGTACTGTTC 81133
QY 522 LysSerArgGlnLysArgPheAspProSerAlaSerLeuAspPheLeuTrpAlaCysIle 541
Db 81132 AAGAAATTTCCATGAACAGTACGACTAGCCGCTCTCAACATCTTCGAGCACTGAC 81073
QY 542 HisValProArgIleTrpGlnGlyArgAspGlnArgThrProGlnLysArgGlyGlnLys 561
Db 81072 ACCAATCCGAGCTGTGGAAAGGAGCGACGAAGTACATGTCGAAGAAGTACGCGCGAT 81013
QY 562 LeuValLeuArgValGlnGlyProGluLeuIleSerLeuValGlnLeuLeu----- 578
Db 81012 GCCTTCTTCATGCTGAGAGACTCTGAACCTGAGACCGCTTCCACATTCATCTTCAGAA 80953
QY 579 ---LeuAlaGlnAlaGlnTrpArgSerGlnAspGlyAspThrAlaAlaCysSerLeuIle 597
Db 80952 GGGCTTTCCGAGGTCAAGCTGACAGCACTAATGATTCAAGCTTCTGCTG----- 80899

QY 598 GlnAlaArgLeuProLeuLeuLeuSerCysCysGlyAspArgLeuSerValArgLys 617
 Db 80898 -----CGAATGAACTACTGTTTAACTCAGAGAGAAAGACAGACATCATGTCAG 80845
 QY 618 ValThrLeuHisLeuSerGlyCysIleGlnGlnTrpGlyLysPheValLeuGlyArgArg 637
 Db 80844 GTATGAGACAGCTGGAG-----AAAAGTTCGGTCCGACTACTTG 80803
 QY 638 CysArgAspLeuLeuLeuGlnLeuTrpLeuGlnArgProGluLeuArgValProValPro 657
 Db 80802 AAACGTGAGGTCTCCAGAGATGATCATGATATCCGGCAGATTTGACAAACCG 80743
 QY 658 -----GluValLeuLeuHisSerGlyAlaAlaSerSer 670
 Db 80742 GGCAAAACCGCGAGAGCGGTACAAATTCAGAACTGAGAGCG----- 80698
 QY 671 ValCysArgLeuAspGlyLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAsp 690
 Db 80697 -----TGGCAGCGGACAAAGGTGTCCACAAATCTGATCACTGCTGGCAGTCTGGCC 80641
 QY 691 SerArgAlaLeuGlnAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuAlaVal 710
 Db 80640 AAAAAGACCTTGTAAACCTGTCTCCAGACAGAGCTGCTTCCAGAGCTGCTGCC 80581
 QY 711 AlaHisProLeuLeuLeuLeuArgHisLeuProMetIleAlaAlaLeuLeuHisGlyArg 730
 Db 80580 TCCACCCCGTCTCTCTCTCTCTGCGACCTGCGCTGCTGCTGCTATTCAGAGCTGCG 80521
 QY 731 ThrHisLeuAsnPheGlnGlnGlnArgGlnGlnAsnHisLeuSerCysPheLeuHisVal 750
 Db 80520 GCCCAGCTCAGCAATGAAGGCCCTCGCGGAGAGACACACTTCCACGATTTGTACAGATC 80461
 QY 751 LeuGlyLeuLeuGlnLeuLeuGlnProHisValPhe-----ArgSerGlnHis 766
 Db 80460 CTGAGGAGCTGGAGCTGTGCAACCACTTGTAGGAGCGGTACAGAACAGATC 80401
 QY 767 GlnGlyAlaLeuTrpArgCysLeuLeuSerPheIleArgLeuLeuLeuAsnTrpArgLys 786
 Db 80400 CAAAACAGCTG-----TCGTCTACTTCACTCTTCAAG-----CACACAGC 80356
 QY 787 SerSerArgHisLeuAlaAlaPheIleAsnLysPheValGlnPheIleHisLysTrpIle 806
 Db 80355 AACGTAAAGAGCGCTCCCAATCTGACACAGATTTGTGCGAGTCTCCAGCTTACATC 80296
 QY 807 ThrTyrsAlaProAlaAlaIleSerPheLeuGlnLysHisAlaAspProLeuHisAsp 826
 Db 80295 AACGTACATCCCTCAAGTCTCTCTCTCATGACAGTACGCGCATTCCTGAAGAA 80236
 QY 827 LeuSerPheAsnSerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeu 846
 Db 80235 CTTCACCCCAAGTACACCTCTACTAGGCAAACTGCAAGGTTTGGTCCAGGCCGTCCTG 80176
 QY 847 ProSerArgAspArgPheThrAspArgGlyLeuAspArgGlyLeuSerLeu 861
 Db 80175 CTGACAGCAAGTGCAGCTGCGCAGCAAGTGTGACACAGCAAGATTAACTACAGTAC 80116
 QY 861 ----- 861
 Db 80115 GATCTGATGAGCATTTGATGTAAAGCCATCGGCGCAGAACCCGCTGTAAAGAGAT 80056
 QY 862 -----GluGlnLysSerSerAlaGlySerLeuPro 871
 Db 80055 CCCATCGAAGTGAATCCGCAAAACCCATGATCCAAAGCAGACATAGGCTTATTCG 79996
 QY 872 LeuValSerValSerLeuPheThrProLeuThrAlaAlaGluMetAlaProTyr----- 889
 Db 79995 GTTCTAACCTTTGGCTGTACAGCGCGTGCAGACTACACGCGCATATCGCGCACTTCTC 79936
 QY 890 -----MetLysArgLeuSerArgGlyGlnThrValGlnAspLeuLeuGlnValLeuSer 907
 Db 79935 GATCTGTCAAGATCAAAAGCAAGTCAACACGAGAGACGTGTCTTGGGTCCATCCAG 79976

QY 908 AspIleAspLeuSerArgArgArgProGluIleLeuSerPhePheSerThrAsnLeu 927
 Db 79875 GAGCTTGAGAGCCCTCACTTCCAGAGATTTGTCTTCAACAGAGCTTTCGAACAGCTG 79816
 QY 928 GlnArgLeuMetSerSerAlaGluGlnCysCysArgAsnLeuAlaPheSerLeuAlaLeu 947
 Db 79815 CTCACCTTATATTCTGCGGAGTGCACAGATCCGAGTCCATCGCTTCATCATCTGATC 79756
 QY 948 ArgSerMetGlnAsnSerPro-----SerIleAlaAlaPheLeuProThrPhe 964
 Db 79755 AGGATCTGAAGACACATCCCGGCAATTCGAGACATCAACTGTGACCCCTTACAGCTTAC 79696
 QY 965 MetLysCysLeuGlySerGlnAspPheGlyValValGlnThrAlaLeuArgAsnLeuPro 984
 Db 79695 ATTACAGCTCTCGCGAGCAAGAACTCTCTGTCGACCGACCGCATTCGATTCGCG 79636
 QY 985 GluTrpAlaLeuLeuCysGlnGlnIleAlaAlaValLeuLeuHisArgAlaPheLeuVal 1004
 Db 79635 GAGATGTCGCTGCTGTCGACAGAAACGACATGCATCTTACAGTGGCTTCTGCTG 79576
 QY 1005 GlyMetLysGlyGlnMetAspProSerArgGlnIleSerGlnAlaLeuArgIleLeu 1023
 Db 79575 GGCTTGAAGTGTGCTGACACTGACACCTGACAGATTTGAAAGGTTCACAGACTCTGA 79519

RESULT 12
 AC004642/c 148432 bp DNA linear INV 01-MAY-1998

LOCUS
 DEFINITION
 ACCESSION
 AC004642 AC004302 AC003551 AC003552 AC003553 AC003554 AC003555
 AC003556 AC003557 AC003558 AC004308 AC003575 AC003848 AC003849
 AC003576 AC003577 AC003578 AC003579 AC003580 AC003581 AC003582
 AC003583 AC003584 AC003585 AC003586 AC003587 AC003588
 AC003589 AC003590 AC003591 AC003592

VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Drosophila melanogaster (Subclones in Sac from Pl clones DS00543 (D193) and DS02867 (D200)) DNA.

Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
 AUTHORS
 Gelink, S.E., Aghavani, A., Arcina, T.T., Baxter, E., Blazek, R.G.,
 Chew, M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A.,
 Hummasti, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B.,
 Lomocan, M.A., Mazda, P., Mok, M.S., Mostreli, A.R., Mostreli, M.,
 Nixon, K., Paclebo, J.M., Park, S., Pfeiffer, B., Punch, D., Snit, E.,
 Twomey, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.L.
 and Kimmel, B.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Sequencing of Drosophila chromosome 2R, region 60B1-60B10
 unpublished (1997)
 2 (bases 1 to 148432)

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Submitted (01-MAY-1998) Berkeley Drosophila Genome Project, MS
 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, US
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory, MS 64-121
 Berkeley, CA 94720

COMMENT
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our sequence
 archive Web site
 (http://fruitfly.berkeley.edu/sequence-archive.html) or send email
 to drosophil@imngc.lbl.gov.

OY	598	glnAlaIaIeProLeuLeuLeuSerGlyScyScySglYAspAspGluSerValArgLys	617
Db	48894	-----CGAATGACCTTACTGTTTAAGCTCACAGAGAAAGCAGACACTCTATGTCAG	48841
OY	618	ValThrGluHisLeuSerGlyCysIleGlnIleTrpGlyAspSerValLeuGlyArg	637
Db	48840	GTAATGGAGCAGCTGGAG-----AAAAGTCCGGTCCGACACTCTG	48799
OY	638	CysArgAspIleuLeuGlnIleuYrLeuGlnArgProGluLeuArgValProValPro	657
Db	48798	AAACTGACAGGTCTCCAGCAGATGTACATCATGTATCCGGCATAAAGTTCTGAACCG	48739
OY	658	-----GluValLeuHisSerGluGlyAlaAlaSerSer	670
Db	48738	GGCAAAACCGGCGAGCAGGCGTAAACAATTTGCAGATCTGAAGGC-----	48694
OY	671	ValCysLysLeuAspGlyLeuIleHisArgPheIleThrLeuLeuAlaAspThrSerAsp	690
Db	48693	--TCCACGGCGGACAGAGGTGTCCACCAATCGATCGATCACTGCTGGCGAGCTGGCGC	48637
OY	691	SerArgAlaLeuGluAlaAsnArgGlyAlaAspAlaSerMetAlaCysArgLysLeuVal	710
Db	48636	AAAAAGACTTTGAAACCTGTCCACGGACACCGAGCTGTCTTCGCAAGCTGGCTGCC	48577
OY	711	AlaHisProLeuLeuLeuLeuArgHisLeuProMetIleAlaIleLeuHisGlyArg	730
Db	48576	TCCACCGCGCTCTCTCTCTCCCTGCGCAGGTGGCGTGGCTGTCCATATGACAGAGTGGC	48517
OY	731	ThrHisLeuAsnPheGlnGluPheArgGlnAlaAsnHisLeuSerCysPheLeuHisVal	750
Db	48516	GCCACGCTCAGATGAAAGCCCTGGCGGAGGAGCACCCACTTCACCGATTTGTACAGATC	48457
OY	751	LeuGlyLeuLeuGluLeuLeuGlnIleProHisValPhe-----ArgSerGluHis	766
Db	48456	CTAGACGAGCTGTGACTGTCTGTGCACCAACCACTTTGAGAGAGCGGTACAGAACAGATC	48397
OY	767	GlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeuLeuLeuAsnTyrArgLys	786
Db	48396	CAAAACAGCGTCTGCTGTCTCACTTCAACCTTCTCAAG-----CACCAACAC	48352
OY	787	SerSerArgHisLeuAlaAlaPheIleLeuHisLysPheValGlnPheIleHisLysTyrIle	806
Db	48351	AAACGTAAGAGAGCGCTGCCAAATGCTGCACAAAGTTGTGTGCACATGCTCCAGCCTATATC	48292
OY	807	ThrTyrAsnAlaProAlaAlaIleSerPheLeuGlnLysHisAlaAspProLeuHisAsp	826
Db	48291	AACTCATATCCCTCAAGTGCCTGCTCTTCATCGCAACAGTACGTGGCGCATCTGAAGGA	48232
OY	827	LeuSerPheAspAsnSerAspLeuValMetLeuLysSerLeuLeuAlaGlyLeuSerLeu	846
Db	48231	CTTGGAGCCAGTACACCTCTACTAGCAAACTGCAGAGTTTGTGTCCAGGCCGTCCCTCG	48172
OY	847	ProSerArgAspAspArgThrAspArgLeuAspArgGluGlnGly-----	861
Db	48171	CTGCAGACAGTGTCACTCGGCGAGCAGGAATTTGGACGAGAGGAAGTTAAGTACAGATAC	48112
OY	861	-----	861
Db	48111	GATCTGATGACATTTGGATGTAAAGCCATCGGCCACCAACCCCGTGTAAACAGAGAT	48052
OY	862	-----GluGlnGluSerSerAlaGlySerLeuPro	871
Db	48051	CCCATGGAAGTAAATCCGCAAAACACCATGCATCCACAGCAGCATAGTAGGGTCCCTTATCG	47992
OY	872	LeuValSerValSerLeuPheThrProLeuThrAlaIleGluMetAlaProTyr-----	889
Db	47991	GTCTTAACCTTGGCTCTACAGCCGGTCCGAACCTACACGAGACATATCCGCCACATTCCTC	47932
OY	890	-----MetLysArgLeuSerArgLysGlnThrValGluAspLeuLeuGluValLeuSer	907
Db	47931	GATCTGGTCAATCATTAAGAGATGCCAACGAGAGAGCGTGTCTTGGTCCCATTCGAG	47872

Oy	908	AspIleAspGluMetSerArgArgArgProGluIleuSerPhePheSerThrAsnLeu	927
Db	47871	GAGTGGAGTCCCTCACTTCACAGAGATTGTGTTCATCAACGAGCGTGTGACGACTG	47812
Oy	928	GlnArgIleuMetSerSerAlaGluGluIuGlySerArgAsnLeuAlaPheSerLeuAlaLeu	947
Db	47811	CTCAACCTTAATATTCCTGGCCGAGTGCCTCCACAGATCCGGTCCATCGGCTTCATCATCTGTGATC	47752
Oy	948	ArgSerMetGlnAsnSerPro-----SerIleAlaAlaAlaPheLeuProThrPhe	964
Db	47751	AGGCATCTGAAGCAACAATCCCGGCAATTCCGACATCAACCTGTGCACCCCTTAACGCCCTAC	47692
Oy	965	MetIlyrGlySerLeuGlySerGlnAsnPheLeuValValGlnThrAlaAlaLeuArgAsnLeuPro	984
Db	47691	ATTTCAGTCCCTCGCGGAGCAAGACTCCTCGGTGGCGACGCGAGCGCCATTGACACATTCGGC	47632
Oy	985	GluTyrAlaLeuLeuLeuCysGlnGluIleHisIleAlaIleValLeuLeuHisArgAlaPheLeuVal	1004
Db	47631	GAGATGTCGGTGGTCGTCTGCAGAAACGCCAATCGACATCTTAACGGTGGCCCTTCCTGCTG	47572
Oy	1005	GlyMetIlyrGlyGlnMetAspProSerAlaGlnIleSerGlnValAlaLeuArgIleLeu	1023
Db	47571	GGCTTGAAGTCCGTCTTAACACTGGCCACCAAGATTGAAGAAGTTCTCCAGACTCTTA	47515

[illegible]

Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
 Nusskern,D.R., Pacleb,J.M., Palazolo,M., Piltman,G.S., Pan,S.,
 Pollard,J., Puri,Y., Reese,M.G., Reinert,K., Remington,K.,
 Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
 Stimpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
 Stapleton,M., Strong,R., Sun,E., Sytkas,R., Tector,C., Turner,R.,
 Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Massaman,D.A.,
 Weinstein,G.M., Weissbach,J., Williams,S.M., Woodage,T.,
 Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
 Zaveri,U.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
 Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
 Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
 The genome sequence of *Drosophila melanogaster*
 Science 287 (5461), 2185-2195 (2000)

JOURNAL MEDLINE
 20196006
 10731132
 2 (bases 1 to 303823)
 Amandlides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D.,
 Banzon,J., Beeson,K.Y., Busan,D.A., Carlson,J.W., Center,A.,
 Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,Y.,
 Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferrera,S., Frise,E.,
 Galles,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J.,
 Hoskins,R.A., Hostin,D., Howland,T.J., Idegawa,C., Jalali,M.,
 Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M.,
 Murphy,B., Nelson,C., Nelson,K.A., Nunco,J., Pacleb,J., Paragas,V.,
 Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Piltman,G.S.,
 Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R.,
 Sytkas,R., Tector,C., Tyler,D., Williams,S.M., Zaveri,J.S.,
 Smith,H.O., Venter,J.C. and Rubin,G.M.
 Sequencing of *Drosophila melanogaster* genome
 Unpublished
 3 (bases 1 to 303823)
 Misra,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L.,
 Campelli,K., Hradscky,P., Huang,Y., Kaminker,J.S., Prochnik,S.E.,
 Smith,C.D., Tupy,J.L., Bergman,C.M., Bertman,B.P., Carlson,W.,
 Celinker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Frise,E., de
 Grey,A.D.N.J., Harris,N.L., Krommiller,B., Marshall,B.,
 Millburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E.,
 Shu,S., Smutniak,F., Whitfield,E.J., Yamada,C., Ashburner,M.,
 Gelbart,W.M., Rubin,G.M., Mungall,C.J. and Lewis,S.E.
 Annotation of *Drosophila melanogaster* genome
 Unpublished
 4 (bases 1 to 303823)
 Adams,M.D., Celinker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
 Direct Submission
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 5 (bases 1 to 303823)
 Lewis,S.E.
 Direct Submission
 Submitted (31-MAY-2002) University of California Berkeley, 539 Life
 Sciences Addition, Berkeley, CA 94720, USA
 On Jun 28, 2002 this sequence version replaced gi:7291637.
 location/Qualifiers
 1. 303823
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /chromosome="2"
 /map="2R"
 complement(11873..23611)
 /gene="CG11169"
 /map="60A9-60A10"
 /db_xref="FLYBASE:FBan0011169"
 /db_xref="FLYBASE:FBgn0034905"
 complement(join(11873..12379,18020..18143,23421..23611))
 /gene="CG11169"
 /product="CG11169-RA"
 /db_xref="FLYBASE:FBan0011169"
 /db_xref="FLYBASE:FBgn0034905"
 complement(join(11873..12379,18020..18091))
 /gene="CG11169"
 /note="CG11169 gene product"
 /codon_start=1

CDS
 1
 /codon_start=1

/product="CG11169-PA"
 /protein_id="AA047060.1"
 /db_xref="GI:7291638"
 /db_xref="FLYBASE:FBan0011169"
 /db_xref="FLYBASE:FBgn0034905"
 /translation="MSSPLVLFALLVSTICQGLPDLKQMTPEQGLSVFHYDTHD
 AVHYELVQIHHNNHNNRRRRSGKPKYNAALPHHHVKRDKSNAYSELKHEALA
 SGGHLELAPPEVSAIKSHNVSFSAFGQHLNLRLATQGLFKGAPHLQMLMTVGSPPNAT
 HGIDLEIVHEQVSRVYSMTOLKIKHKVY"
 complement(16742..17216)
 /gene="CG15800"
 /map="60A10-60A10"
 /db_xref="FLYBASE:FBan0015800"
 /db_xref="FLYBASE:FBgn0034904"
 complement(16743..17216)
 /gene="CG15800"
 /product="CG15800-RA"
 /db_xref="FLYBASE:FBan0015800"
 /db_xref="FLYBASE:FBgn0034904"
 complement(16743..17216)
 /gene="CG15800"
 /note="CG15800 gene product"
 /codon_start=1
 /product="CG15800-PA"
 /protein_id="AA047061.1"
 /db_xref="GI:7291639"
 /db_xref="FLYBASE:FBan0015800"
 /db_xref="FLYBASE:FBgn0034904"
 /translation="WRFETRDGRTISVNTALLESLSIVREMKRGVPEDEPVPL
 HGHSNVLLKILMAEYHERKEPAAWDSKDPLEIEIQLISWDEFLKVELDISIC
 KIMECNVLDIPMLYKCAQKLYFLSSREPEYRGYVEKIPRODHFIRDOKE"
 27566..28291
 /gene="CG30177"
 /map="60A10-60A10"
 /db_xref="FLYBASE:FBan0030177"
 /db_xref="FLYBASE:FBgn0050177"
 27566..28291
 /gene="CG30177"
 /product="CG30177-RA"
 /db_xref="FLYBASE:FBan0030177"
 /db_xref="FLYBASE:FBgn0050177"
 27671..28291
 /gene="CG30177"
 /note="CG30177 gene product"
 /codon_start=1
 /product="CG30177-PA"
 /protein_id="AA068270.1"
 /db_xref="GI:2162663"
 /db_xref="FLYBASE:FBan0030177"
 /db_xref="FLYBASE:FBgn0050177"
 /translation="MLRLVYIYGTLLPGWSLRAIGHDRHFLEMLKVVYVALLO
 GLGLVTPFLGLGSFYAGCKILISVGLMFSAPYSTNNHFNLLNTYSMGKICPVYKCV
 RWHRKVCNRIKRGFMKSLVTYVMPRODEHITPERLRNNRLKLEKSNLMTQIAKNSA
 DMRHQHSPMIRHSPRLRAELDKFNESSMHNVIDYLGQGGDQO"
 complement(30339..30720)
 /gene="CG13561"
 /map="60A11-60A11"
 /db_xref="FLYBASE:FBan0013561"
 /db_xref="FLYBASE:FBgn0034906"
 complement(join(30339..30494,30556..30720))
 /gene="CG13561"
 /product="CG13561-RA"
 /db_xref="FLYBASE:FBan0013561"
 /db_xref="FLYBASE:FBgn0034906"
 complement(join(30339..30494,30556..30720))
 /gene="CG13561"
 /note="CG13561 gene product"
 /codon_start=1
 /product="CG13561-PA"
 /protein_id="AA047062.1"
 /db_xref="GI:7291640"
 /db_xref="FLYBASE:FBan0013561"
 /db_xref="FLYBASE:FBgn0034906"
 /db_xref="GI:7291640"

CDS
 1
 /codon_start=1

Oy	579	-----leuallagluagluuthrargserglinasprolysphrhlalaacysserleulle	597
Db	268908	GGCGTTTCCGAGGTCAAGAGCTGGACAGCAAAAACATCATTTTCAAGCTCTCG-----	268961
Oy	598	glinalaargleuproleuleuleuSerCysCysgslaspsapgluserValaIatlys	617
Db	268962	-----CGAATGAACTACTTATTAACTCAACAGAGAAAGGACGAGACCTCATCTGTCAG	269015
Oy	618	ValThrghlnHisleuSerGlyCysIleaglInlTrpPolyspserValleuGlyArgArg	637
Db	269016	GTAAATGGAGCGACGTGGAG-----AAAAGTTTCGGGTCCGACTCTTG	269057
Oy	638	CysArgAspLeuLeuLeuGlnLeuTyrLeuGlnArgProGluIleArgValProValPro	657
Db	269058	AAATCGCAGGTCCTCCACAGATATACATCATGTATCCGGCATCAAGTTTCTGAACCG	269117
Oy	658	-----GluValleuLeuHisSerGluGlyAlaAlaIaSerSer	670
Db	269118	GGCAAAACCGCGACGAGCGCTCAAAATTCAGCAATTCGAATTCGAAGGC-----	269162
Oy	671	ValCysIysleuAspGlyLeuIleHisArgPheIleThrleuLeuAlaAspTrpSerAsp	690
Db	269163	--TCCACGGCGGACAAAGGTCTCCAAACAATCTGATCACTGGCTTGGCGCACTGGTGGC	269219
Oy	691	SerArgAlaLeuGluAsnArgGlyAlaAlaSerMetAlaCysArgIysleuAlaVal	710
Db	269220	AAAAAGACTTTGAAACCCCTTCCACGGACACGAGCTGCTCTCCACAGCTGGCTGCC	269279
Oy	711	AlaHisProleuLeuLeuArgHisleuProMetIleAlaAlaLeuLeuHisGlyArg	730
Db	269280	TCCCAACCGGCTTCTTCTTCCCTGGCACCTGGCGCTGCTGTCATTAAGCAAGGTGG	269339
Oy	731	ThrHisleuAsnPheGlnGluPheArgGlnGlnAsnHisleuSerCysPheLeuHisVal	750
Db	269340	GCCCAAGTCACACATGAAGGCCCTTCCAGGAGACACCACTTCCACGATTTGTACAGATC	269399
Oy	751	LeuGlyLeuLeuGluLeuLeuGlnProHisValPhe-----ArgSerGluHis	766
Db	269400	CTGAGGAGCGCTGGAGCTGCTCTCAACAACATCTTTGACAGAGCGGTCAAGAACAGATC	269459
Oy	767	GlnGlyAlaLeuTrpAspCysLeuLeuSerPheIleArgLeuLeuLeuAsnIlyArgIlys	786
Db	269460	CAAAACACGCTG---TCTGCTTACTTCAACTCTCTCAAG-----CACACAGC	269504
Oy	787	SerSerArgHisleuAlaAlaPheIleAsnIysPheValGlnPheIleHisIlysIyrIle	806
Db	269505	AACGTAAGAAGCGCTGCCAANAATCTGACAAAGATTGTGCGAGATCTCCAAAGCTACATC	269564
Oy	807	ThrIyrAsnAlaProIleAlaIleSerPheLeuGlnIlyHisAlaAspProleuHisAsp	826
Db	269565	AACATCAATCCCTCAAGGCTCTCTCTTCATTCGAACAGTACGTGGGCATCTCGAAGAA	269624
Oy	827	LeuSerPheAspAsnSerAspLeuValMetLeuIysSerLeuLeuAlaGlyLeuSerLeu	846
Db	269625	CTTGAGCGCAGTACACCTCACTCAAGCAACTGCGAGGTTTGGTCCAGGCGGTTGGCCTG	269684
Oy	847	ProSerArgAspAspArgThrAspArgGlyLeuAspGluGluGly-----	861
Db	269685	CTGCAGCACAAGTGCACCTCGCGACAGGAATTGGACGACGACGAAAGTTAAGTACGAGTAC	269744
Oy	861	-----	861
Db	269745	GATCGCATGACATTTTCGATGTAAAGCATCGGCACGAAGCCCGTGTACACAGAGAT	269804
Oy	862	-----GluGlnGluSerSerAlaGlySerLeuPro	871
Db	269805	CCCATCGAATGGAATCCGCAACAACCCATCGATTCAGACGAGCTAGAGGGTCCCTTATCG	269864
Oy	872	LeuValSerValSerLeuPheThrProleuThrAlaIaGluMetAlaProIy-----	889
Db	269865	GTTCAACCTTGGGCTGCTACAGCGGCTCGAATCAACGAGCATATCTGCCGCACTTCTCTC	269924

TITLE	JOURNAL
2 (bases 1 to 144097)	Worley, K.C.
Direct Submission	Submitted (05-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 144097)	Worley, K.C.
Direct Submission	Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 18, 2002 this sequence version replaced gi:20452937.	
Genome Center	
Center: Baylor College of Medicine	
Center code: BCM	
Web site: http://www.hgsc.bcm.tmc.edu/	
Contact: hgsc-help@bcm.tmc.edu	
Project Information	
Center project name: GNMW	
Center clone name: CH230-101110	
Summary Statistics	
Sequencing vector: Plasmid:	
Chemistry: Dye-terminator Big Dye: 100% of reads	
Assembly program: Phrap; version 0.990329	
Consensus quality: 62007 bases at least Q40	
Consensus quality: 68326 bases at least Q30	
Consensus quality: 73396 bases at least Q20	
NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)	
NOTE: This is a 'working draft' sequence. It currently consists of 72 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	
1	1060: contig of 1060 bp in length
1061	1160: gap of unknown length
1161	2270: contig of 1110 bp in length
1271	2370: gap of unknown length
2371	3371: contig of 1001 bp in length
3372	3471: gap of unknown length
3472	4834: contig of 1363 bp in length
4835	4934: gap of unknown length
4935	5992: contig of 1058 bp in length
5993	6092: gap of unknown length
6093	7670: contig of 1578 bp in length
7671	7770: gap of unknown length
7771	9212: contig of 1442 bp in length
9213	9312: gap of unknown length
9313	10541: contig of 1229 bp in length
10542	10641: gap of unknown length
10642	12072: contig of 1431 bp in length
12073	12172: gap of unknown length
12173	13552: contig of 1380 bp in length
13553	13652: gap of unknown length
13653	15096: contig of 1444 bp in length
15097	15196: gap of unknown length
15197	16222: contig of 1026 bp in length

